tijdschrift van het

nederlands elektronica- en radiogenootschap

deel 41 - nr. 3 - 1976

nederlands elektronica- en radiogenootschap

Nederlands Elektronica- en Radiogenootschap
Postbus 39, Leidschendam. Gironummer 94746 t.n.v.
Penningmeester NERG, Leidschendam.

HET GENOOTSCHAP

Het Genootschap stelt zich ten doel in Nederland en de Overzeese Rijksdelen de wetenschappelijke ontwikkeling en de toepassing van de elektronica en de radio in de ruimste zin te bevorderen.

Bestuur

Lidmaatschap

Voor lidmaatschap wende men zich tot de secretaris. Het lidmaatschap staat -behoudens ballotage- open voor academisch gegradueerden en hen, wier kennis of ervaring naar het oordeel van het bestuur een vruchtbaar lidmaatschap mogelijk maakt. De contributie bedraagt fl. 45,--.

Studenten aan universiteiten en hogescholen komen bij gevorderde studie in aanmerking voor een junior-lidmaatschap, waarbij 50% reductie wordt verleend op de contributie. Op aanvraag kan deze reductie ook aan anderen worden verleend.

HET TIJDSCHRIFT

Het tijdschrift verschijnt zesmaal per jaar. Opgenomen worden artikelen op het gebied van de elektronica en van de telecommunicatie.

Auteurs die publicatie van hun wetenschappelijk werk in het tijdschrift wensen, wordt verzocht in een vroeg stadium kontakt op te nemen met de voorzitter van de redactie commissie.

De teksten moeten, getypt op door de redactie verstrekte tekstbladen, geheel persklaar voor de offsetdruk worden ingezonden.

De abonnementsprijs van het tijdschrift bedraagt f 45,--. Aan leden wordt het tijdschrift kosteloos toegestuurd.

Tarieven en verdere inlichtingen over advertenties worden op aanvrage verstrekt door de voorzitter van de redactiecommissie.

Redactiecommissie

DE EXAMENS

De examens door het Genootschap ingesteld en afgenomen zijn:

a. op lager technisch niveau:"Elektronica monteur NERG"
b. op middelbaar technisch niveau: Middelbaar Elektronica Technicus NERG"

Brochures waarin de exameneisen en het examenreglement zijn opgenomen kunnen schriftelijk worden aangevraagd bij de Administratie van de Examencommissie.

Voor deelname en inlichtingen wende men zich tot de Administratie van de Examencommissie NERG, Genemuidenstraat 279, den Haag, gironummer 6322 te den Haag.

Examencommissie

EENVOUDIGE VOCODER VOOR DIGITALE SPRAAKOVERDRACHT

R.J. Sluyter

Philips Natuurkundig Laboratorium

This paper describes a vocoder for use in a speech communication system which transmits digital signals at a rate of 4800 bps. The vocoder is based on formant analysis-synthesis. The first formant, however, is transmitted in its original form. The second formant synthesis is done by so-called voice-excitation, and the third formant is noise-excited. This concept leads to a good balance between performance and complexity so that a useful system is obtained.

INLEIDING

Het overdragen van gespreksignalen in digitale vorm over telefoonverbindingen ondervindt een groeiende belangstelling o.a. in verband met geheimhouding van gesprekken. Als het spraaksignaal eenmaal beschikbaar is in digitale vorm dan kan met een z.g. "scrambler" een hoge mate van geheimhouding verkregen worden. Vanwege de beperkte bandbreedte van telefoonkanalen zijn slechts lage bitsnelheden mogelijk. Standaard bitsnelheden voor datatransmissie over telefoonlijnen zijn o.a. 2400 bps, 4800 bps en 9600 bps. Om gespreksignalen te kunnen coderen met deze lage bitsnelheden en met behoud van een zo goed mogelijke spraakkwaliteit, moet men gebruik maken van spraak analyse-synthese systemen (vocoders).

Om de complexiteit van het complete geheimhoudingssysteem binnen redelijke grenzen te houden, is gekozen voor een transmissiesnelheid van 4800 bps. Een transmissiesnelheid van 2400 bps zou een ingewikkelde vocoder vereisen, maar de datatransmissieapparatuur (de modem) kan eenvoudig zijn. Anderzijds zou bij een transmissiesnelheid van 9600 bps volstaan kunnen worden met een eenvoudige vocoder, maar nu heeft men een complexe modem nodig. Een transmissiesnelheid van 4800 bps echter, geeft aanleiding tot een systeem waarbij zowel de vocoder als de modem relatief eenvoudig kunnen zijn.

SPRAAK EN FORMANTEN

Aan de zendzijde van een vocoder wordt het spraaksignaal geanalyseerd, hetgeen een aantal relatief langzaam variërende parametersignalen oplevert. Aan de ontvangzijde wordt het spraaksignaal gesynthetiseerd welke synthese gestuurd wordt door de parametersignalen. Omdat de verschillende soorten spraakklanken op verschillende wijze gesynthetiseerd moeten worden, maken we onderscheid tussen stemhebbende en stemloze klanken. De belangrijkste stemhebbende klanken zijn de klinkers.

Een klinker ontstaat wanneer de mondholte akoestisch wordt aangestoten door luchtdrukimpulsen, welke gegenereerd worden met behulp van de stembanden. De frekwentie waarmee deze impulsen optreden, de grondtoonfrekwentie, kan zo ongeveer tussen 80 en 300 Hz variëren. Naast de grondtoon bevatten de impulsen een groot aantal harmonischen, welke in sterkte afnemen met de frekwentie. De mondholte vertoont in het algemeen een aantal resonanties, waardoor bepaalde delen in het spektrum van de impulsen sterk bevoordeeld worden. Zo'n resonantie wordt een formant genoemd. Fig. 1 geeft een voorbeeld van het spektrum van een klinker. Samenvattende wordt een klinker dus gekarakteriseerd door een grondtoonfrekwentie F_0 en de frekwentieliggingen en relatieve sterkten van de formanten. Deze parameters kunnen slechts langzaam variëren omdat ze bestuurd worden door spierbewegingen. De tijdkonstante waarmee dit gebeurt is enkele tientallen milliseconden groot.

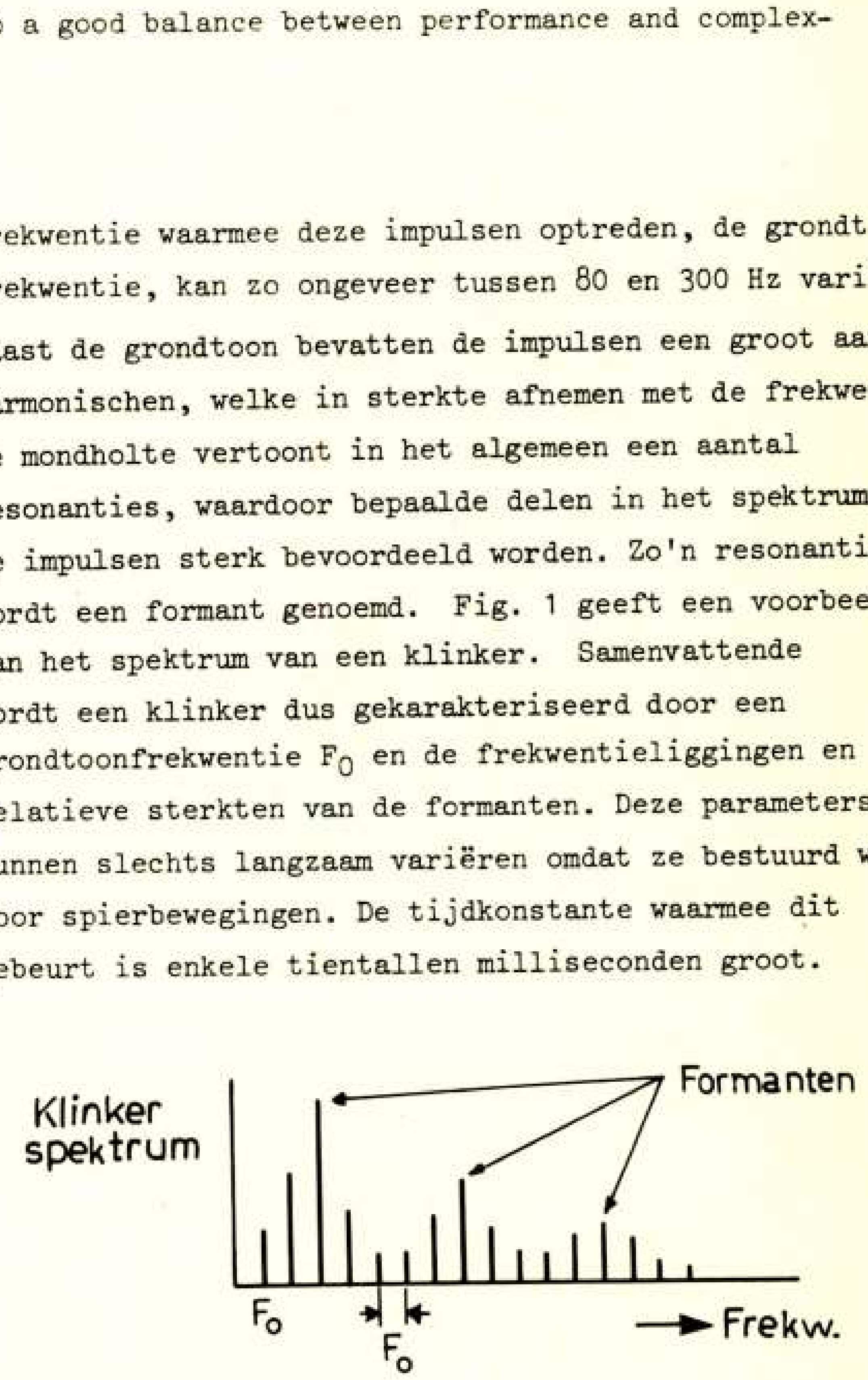


Fig.1 Spektrum van een klinker.

In de band tot 3 à 4 kHz van het spektrum van een klinker treden normaal slechts drie formanten op. Tevens zijn er z.g. formantgebieden aan te wijzen : de eerste formant valt meestal in de band van 300 - 800 Hz, de tweede formant valt in een groot aantal van de gevallen in de band van 800 - 2000 Hz en de derde formant valt meestal in de band van 2000 - 3200 Hz (Flanagan 1965).

In tegenstelling tot het spektrum van een klinker, welke in sterkte afneemt met de frekwentie, is het grootste gedeelte van de energie van stemloze klanken, zoals s en f, vertegenwoordigd in de band boven 2 kHz. Bovendien vertonen deze klanken geen periodiciteit, zoals de klinkers. De stembanden trillen niet en de klank heeft een ruisachtig karakter.

DE VOCODER

Het blokschema van de vocoder is weergegeven in Fig. 2. Het spraaksignaal wordt met behulp van filters opgesplitst in drie kanalen welke ieder een van de formantgebieden omvat. De gebruikte bandbreedtes voor deze formantkanalen zijn 300 - 800 Hz voor het eerste formantkanaal, 800 - 2000 Hz voor het tweede formantkanaal en 2000 - 3200 Hz voor het derde formantkanaal. Aan de ingang van het systeem wordt het spraaksignaal eerst gedifferentieerd waardoor het gemiddelde frekwentiespektrum van dit signaal min of meer vlak wordt. Hierdoor kunnen aan de ingangsfilters dezelfde dempingseisen in het spergebied worden gesteld. In de praktijk blijkt dat een sperdemping van $\geqslant$30 dB voldoende is.

Elk van de formanten wordt na analoge bewerking (analyse) toegevoerd aan een digitaliseringsinrichting. De verkregen digitale signalen worden door een multiplexer samengevoegd tot een synchrone datastroom van 4800 bps. Aan de ontvangzijde wordt deze datastroom aan een demultiplexer toegevoerd, waarna de digitale signalen gedecodeerd worden. De gedecodeerde formantkanalen worden elk na weer analoge bewerking (synthese) aan de uitgang toegevoerd.

Na een verdere toelichting van de gebruikte filters worden de formantkanalen elk apart beschreven wat betreft hun analoge gedeelten. De digitalisering van de verschillende signalen wordt in een aparte paragraaf beschreven.

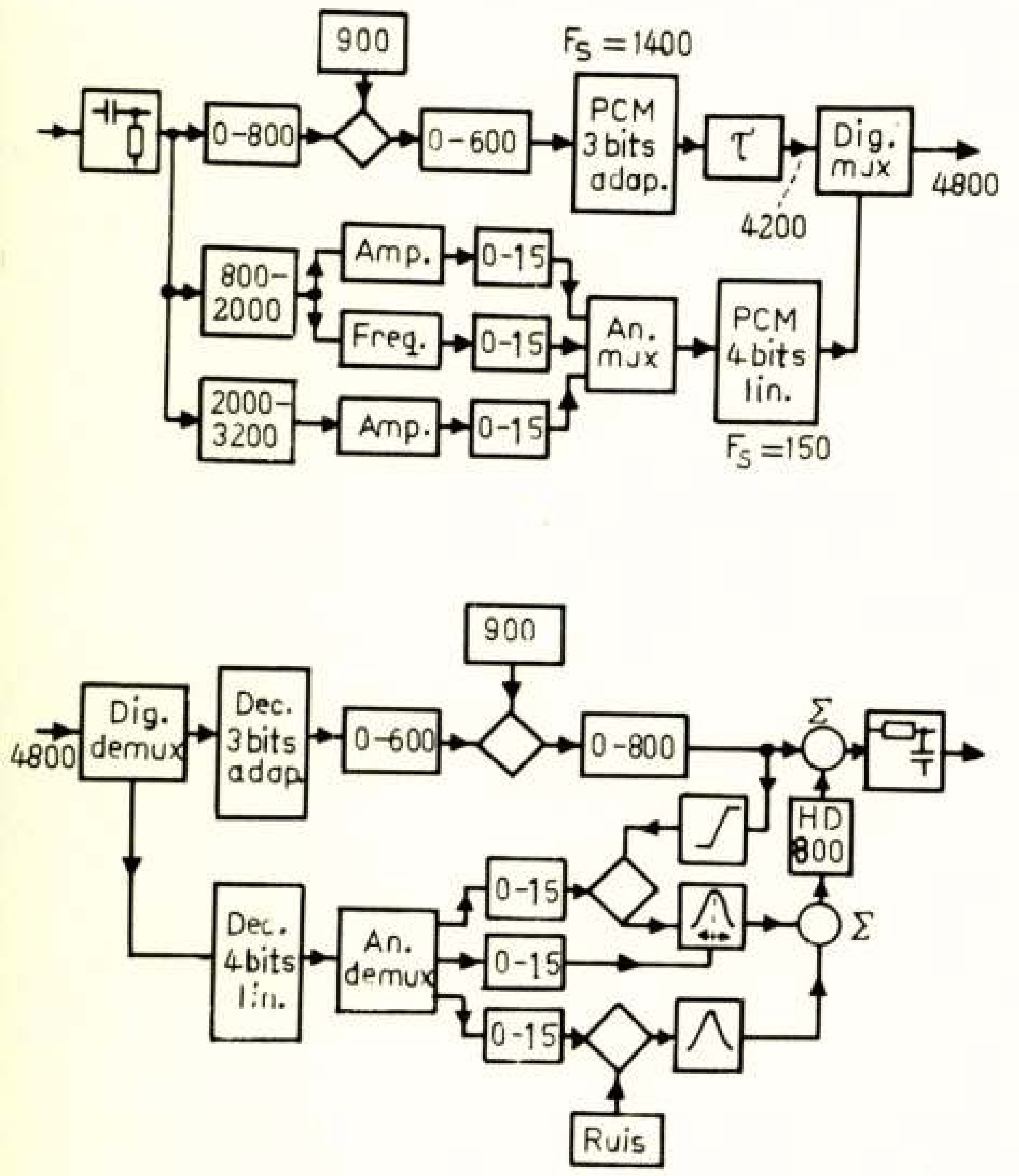


Fig. 2 Blokschema van de vocoder.
Boven : zendgedeelte,
onder : ontvanggedeelte.

Filters

De eisen die aan de filters gesteld worden zijn : een sperdemping van $\geqslant$30 dB, zoals reeds eerder genoemd, en een steile overgang tussen doorlaat en spergebied zodat een goede bandscheiding wordt verkregen. Een geringe rimpel in het doorlaatgebied is geen bezwaar. Tevens streven we naar zo eenvoudig mogelijke filters. Een filtertype dat hieraan goed voldoet is een 4e-orde Cauer filter met een minimale sperdemping van 30 dB en een rimpel in het doorlaatgebied van 1,25 dB. De op de afsnijfrekwentie genormeerde dempingskarakteristiek van dit filter is gegeven in Fig. 3. Naast de laagdoorlaatkarakteristiek is in Fig. 3 ook de hoogdoorlaatkarakteristiek weergegeven. Omdat RC-aktieve filters zijn toegepast, is geen direkte transformatie mogelijk van laagdoorlaat naar banddoorlaat. Daarom zijn de bandfilters elk opgebouwd uit een cascade van een laagdoorlaatfilter en een hoogdoorlaatfilter.

Alle filters in de vocoder zijn van het boven beschreven type, alleen de laagdoorlaatfilters 0 - 15 Hz zijn van een eenvoudiger type. Deze laatste worden in de beschrijving van de formantkanalen nader toegelicht.

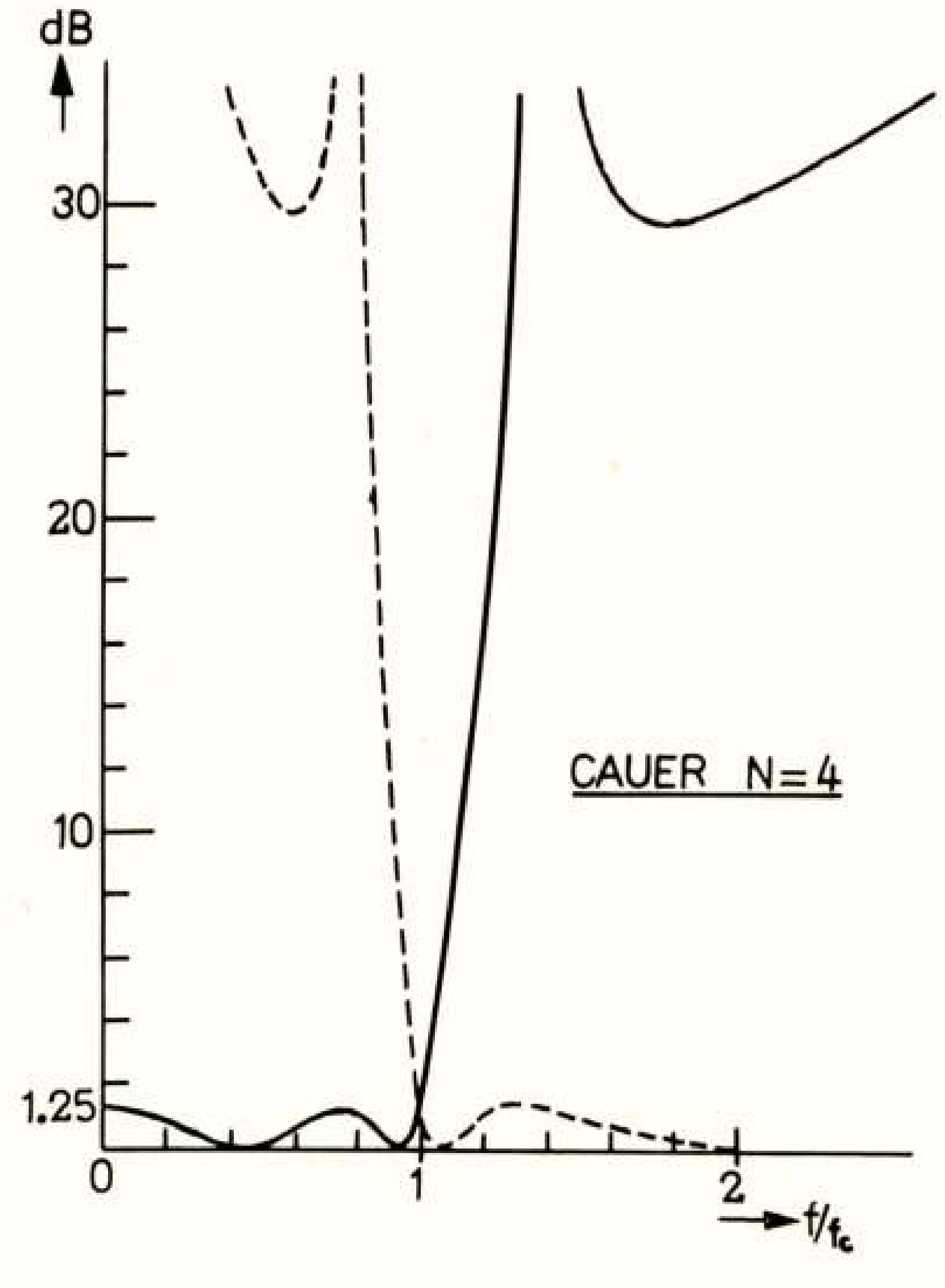


Fig. 3 Genormaliseerde dempingskarakteristieken van de gebruikte filters.

1e- Formantkanaal

Het eerste formantsignaal wordt zonder bandbreedtebeperking overgezonden. Dit heeft twee redenen. Ten eerste verschaft dit een stuk natuurlijke spraak aan de ontvangkant, wat de verstaanbaarheid zowel als de natuurlijk-

heid van het spraaksignaal zeer ten goede komt. Ten tweede bevat dit kanaal grondtooninformatie waarvan met vrucht gebruik gemaakt kan worden voor de natuurgetrouwe synthese van het tweede formantsignaal, waarop wel bandbreedte beperking wordt toegepast.

Het eerste formantkanaal is niet uitgerust met bandfilters maar met laagdoorlaatfilters en modulatoren. Dit vindt zijn oorzaak in de digitalisering van het eerste formantsignaal. De nauwkeurigheid waarmee dit signaal gedigitaliseerd wordt, wordt voornamelijk bepaald door de verhouding van de bitsnelheid van het gedigitaliseerde signaal en de hoogste frekwentie die in het analoge signaal voorkomt. De modulatietrap verschuift met behulp van een draaggolfsignaal van 900 Hz het eerste formantsignaal van de band 300 - 800 Hz naar de band 100 - 600 Hz. De genoemde verhouding wordt hierdoor verbeterd met een niet te versmaden faktor 800/600 = 1,33.

Aan de ontvangstzijde wordt het signaal weer naar de oorspronkelijke band verschoven met een soortgelijke modulatietrap.

2e-Formantkanaal

Bij de analyse van het tweede formantsignaal worden twee grootheden gedetekteerd : de formantamplitude en de formantfrekwentie. Zoals in Fig. 4 is aangegeven, bestaat de amplitudedetektor uit een dubbelfasige gelijkrichter en een laagdoorlaatfilter. Dit laagdoorlaatfilter bestaat uit een enkele RC-sectie met een 3 dB-afsnijfrekwentie van 15 Hz.

De frekwentiedetektor is gebaseerd op het meten van het aantal nuldoorgangen van het signaal per tijdseenheid. Als er een enkele formant in het kanaal aanwezig is, dan is de golfvorm A(t) van het formantsignaal zoals aangegeven in Fig. 4. Het formantsignaal lijkt op het signaal van een door een impuls aangestoten resonantiekring. Dit is volledig in overeenstemming met het gegeven model van de produktie van klinkers. Het formantsignaal A(t) is periodiek met de grondtoonfrekwentie F_0 en de nuldoorgangen worden hoofdzakelijk bepaald door de formantfrekwentie F. Ter verduidelijking is ook het bij A(t) behorende frekwentiespektrum A(f) weergegeven in Fig. 4.

Het circuit van de frekwentiedetektor bevat een clipper welke een tweewaardig signaal afgeeft en zodoende uitsluitend de nuldoorgangen van het formantsignaal handhaaft. Aan de ingang van de clipper wordt een klein signaaltje met een frekwentie van ca. 1400 Hz toegevoegd. Gedurende aanwezigheid van een spraaksignaal is het formantsignaal zo sterk dat dit hulpsignaaltje de tijdstippen van de nuldoorgangen van het formantsignaal nauwelijks beïnvloedt. Gedurende pauzes echter, werkt de clipper op de nuldoorgangen van het hulpsignaaltje zodat de detektor dan toch een gedefinieerde waarde zal af-

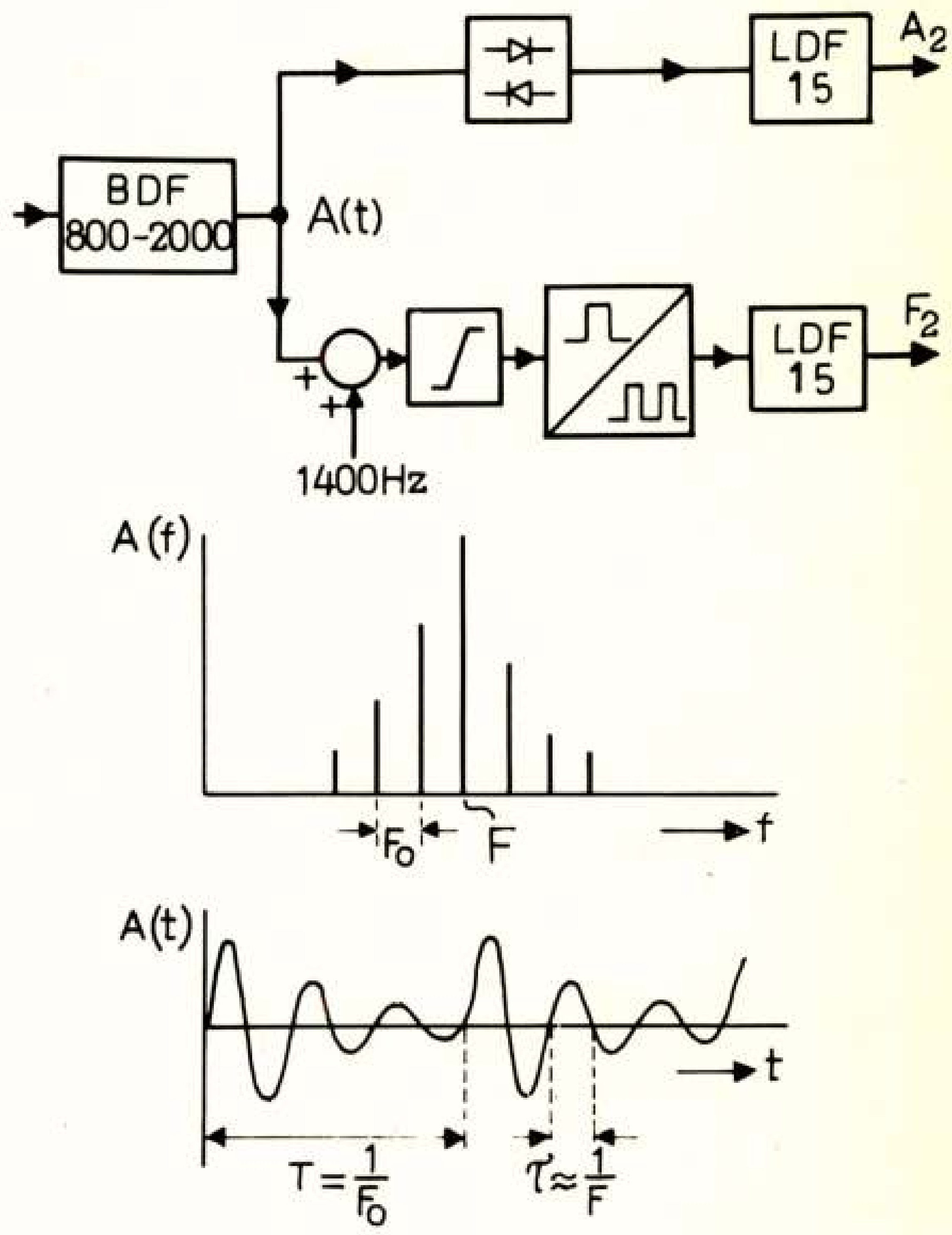


Fig. 4 Boven : Principeschema van de gebruikte formant-analysator.
Onder : Tijd-en frekwentiefunkties van een karakteristiek formantsignaal.

geven. De flanken van het geclipte signaal triggeren een impulsgenerator welke op elke flank een impuls afgeeft van konstante tijdsduur en amplitude. Het laagdoorlaatfilter geeft de gemiddelde waarde van deze impulsreeks af welke dus evenredig is met de formantfrekwentie. Het filter bestaat uit een enkele RC-sectie met een 3 dB-afsnijfrekwentie van 15 Hz.

Op bovenstaande wijze is de bandbreedte van het formantsignaal, welke 1200 Hz omvat, teruggebracht tot enkele tientallen Herzen.

Aan de ontvangzijde wordt het formantsignaal gesynthetiseerd met behulp van een resonantiekring waarvan de resonantiefrekwentie wordt geregeld door de ontvangen frekwentieregelspanning (Flanagan 1960). Deze resonantiekring bootst de eigenschappen van de mondholte voor klinkers na, althans in het tweede formantgebied. Het circuit van de afstembare resonantiekring bestaat uit twee tegengekoppelde operationele versterkers en een variabele versterker met versterking -A. Zie Fig. 5.
De overdrachtsfunktie H(p) van het circuit is :

$$H(p) = \frac{-p\frac{R^2C}{R_0A}}{p^2\frac{R^2C^2}{A} + p\frac{R^2C}{R_0A} + 1} .$$

De resonantiefrekwentie ω_0, de kwaliteitsfaktor Q, de bandbreedte B en de versterking bij ω_0, $H(\omega_0)$, worden gegeven door de betrekkingen welke eveneens in Fig. 5 zijn weergegeven. Hieruit blijkt dat ω_0 gevariëerd kan worden door de versterking A te variëren. Aangezien de ontvangen frekwentieregelspanning F_2 recht evenredig is met de formantfrekwentie, moet A dus kwadratisch geregeld worden :

$$A = kF_2^2,$$

waarin k een passende konstante voorstelt. Nu wordt ω_0 :

$$\omega_o = \frac{\sqrt{A}}{RC} = \frac{\sqrt{k}}{RC} F_2$$

De resonantiefrekwentie ω_0 volgt zodanig de in de zender gemeten formantfrekwentie F. De bandbreedte van de kring is ongeveer 100 Hz gekozen, overeenkomende met de gemiddelde bandbreedte van de tweede formant.

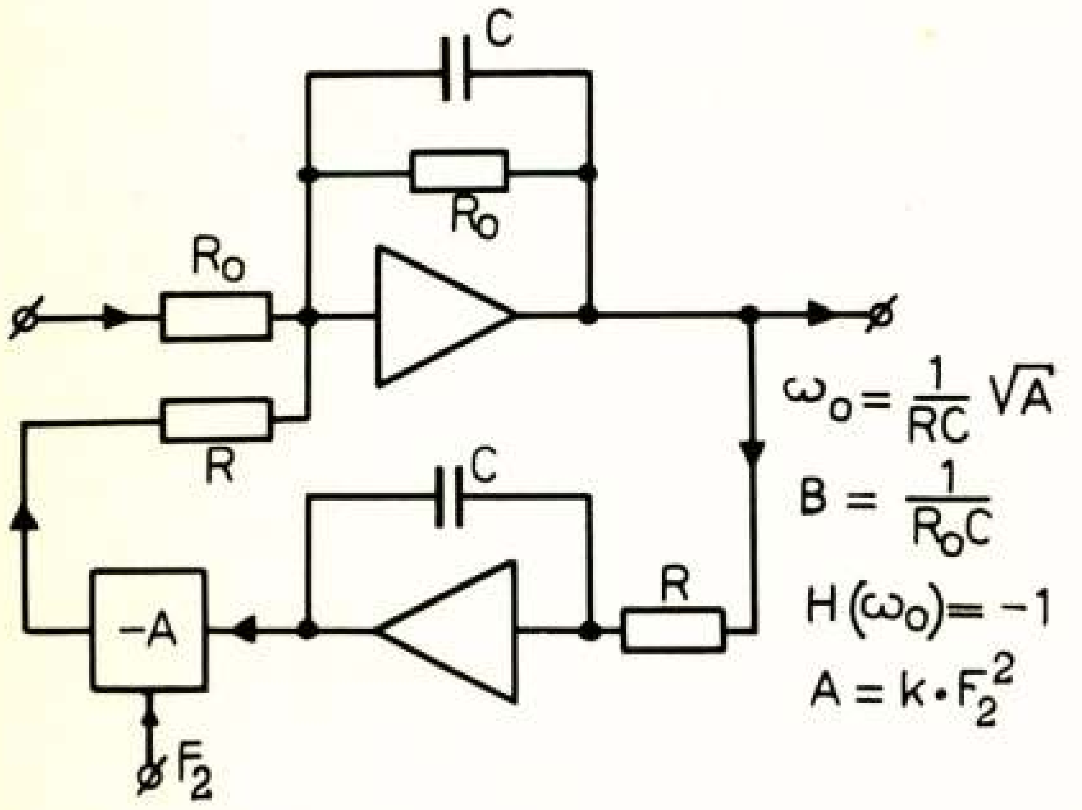


Fig. 5 RC-aktieve realisatie van de afstembare resonantiekring. De afstemming geschiedt door de versterkingsfaktor A te variëren.

De resonantiekring wordt via een amplitudemodulator aangestoten door het vervormde eerste formantsignaal. Deze vervorming wordt bewerkstelligd door een clipper. Het vervormde signaal bevat harmonischen van de grondtoon die aanwezig is in het eerste formantsignaal. Op deze wijze blijft de periodiciteit, of grondtooninformatie, in de tweede formant gehandhaafd. Men kan i.p.v. een clipper ook andere niet-lineaire circuits toepassen zoals bijv. een gelijkrichter, maar een clipper heeft het voordeel dat deze niveau-variaties van het eerste formantsignaal elimineert zodat een gelijkmatiger aanstoten van de resonantiekring verkregen wordt. De juiste niveauvariaties in de tweede formant worden bewerkstelligd door het uitgangssignaal van de clipper te moduleren met de ontvangen amplituderegelspanning.

Het uitgangssignaal van de resonantiekring wordt aan de uitgang van het systeem toegevoerd via een hoogdoorlaatfilter met een afsnijfrekwentie van 800 Hz. Dit filter elimineert eventueel aanwezige signaalcomponenten in het eerste formantgebied.

3e-Formantkanaal

Uit de beschrijving van het tweede formantkanaal blijkt dat dit kanaal uitsluitend is ingericht voor de overdracht van stemhebbende klanken, en in het bijzonder van klinkers. De relatieve bijdrage van de derde formant tot een klinker is slechts gering : gemiddeld ca. -25 dB (Peterson 1952). Het is daarom logisch om het derde formantkanaal niet te gebruiken voor deze kleine bijdrage, maar voor de overdracht van stemloze klanken, die juist in dit frekwentiegebied een groot gedeelte van de energie hebben. Daarom wordt de derde formant ontvanger niet aangestoten door het vervormde eerste formantsignaal, maar door een ruisbron. Door deze maatregel heeft het nagenoeg geen effekt meer om de resonantiefrekwentie van de kring in de derde formant ontvanger te variëren. Hierdoor komt het gehele frekwentiekanaal te vervallen.

De zendzijde van het derde formantkanaal bestaat naast het ingangsfilter nog slechts uit een amplitudedetektor.

De ontvangzijde bestaat uit een vaste resonantiekring welke aangestoten wordt door een met het ontvangen amplituderegelsignaal gemoduleerde ruisbron. De resonantiefrekwentie van de kring is 2500 Hz en de bandbreedte is 200 Hz. Tenslotte wordt het signaal, via het 800 Hz-hoogdoorlaatfilter, aan de uitgang toegevoerd. Het signaal uit de somvormer wordt aan een integrator toegevoerd welke de differentiator aan de ingang van de zender compenseert.

Digitalisering

De digitalisering van de regelsignalen in de tweede en derde formantkanalen gebeurt met lineaire PCM. De signalen worden aan een analoge multiplexer toegevoerd, welke de drie signalen om beurten doorschakelt naar de ingang van de PCM-coder. Deze coder bemonstert zijn ingangssignaal met 150 Hz, zodat de regelsignalen elk met 50 Hz bemonsterd worden. De signaalmonsters worden elk gekwantiseerd met 4 bits, zodat per kanaal een bitssnelheid van 200 bps ontstaat. Het uitgangssignaal van de coder bestaat dus uit een 600 bps datastroom.

Aangezien de beschikbare transmissiecapaciteit 4800 bps is, blijven er nog 4200 bps over voor het digitaliseren van het eerste formantsignaal.

Het codeer-principe dat we gebruiken voor het digitaliseren van het eerst formantsignaal is adaptieve PCM (APCM), (Wilkinson 1971). Het signaal wordt bemonsterd met 1400 Hz en elk monster wordt gekwantiseerd met 3 bits, echter zodanig dat de kwantiseringsniveau's zichzelf aanpassen aan het niveau van het ingangssignaal. Fig. 6 geeft het vereenvoudigde blokschema van de APCM-coder.
De zeven te onderscheiden niveau's worden aan de comparators toegevoerd d.m.v. een weerstandsdeling van de regelspanning V_r. De uitgangsspanningen van de comparators

worden in de binaire converter omgezet in een 3-bits representatie van de aan de regelspanning gerelateerde ingangssignaalwaarde. Het uitgangsregister wordt op de

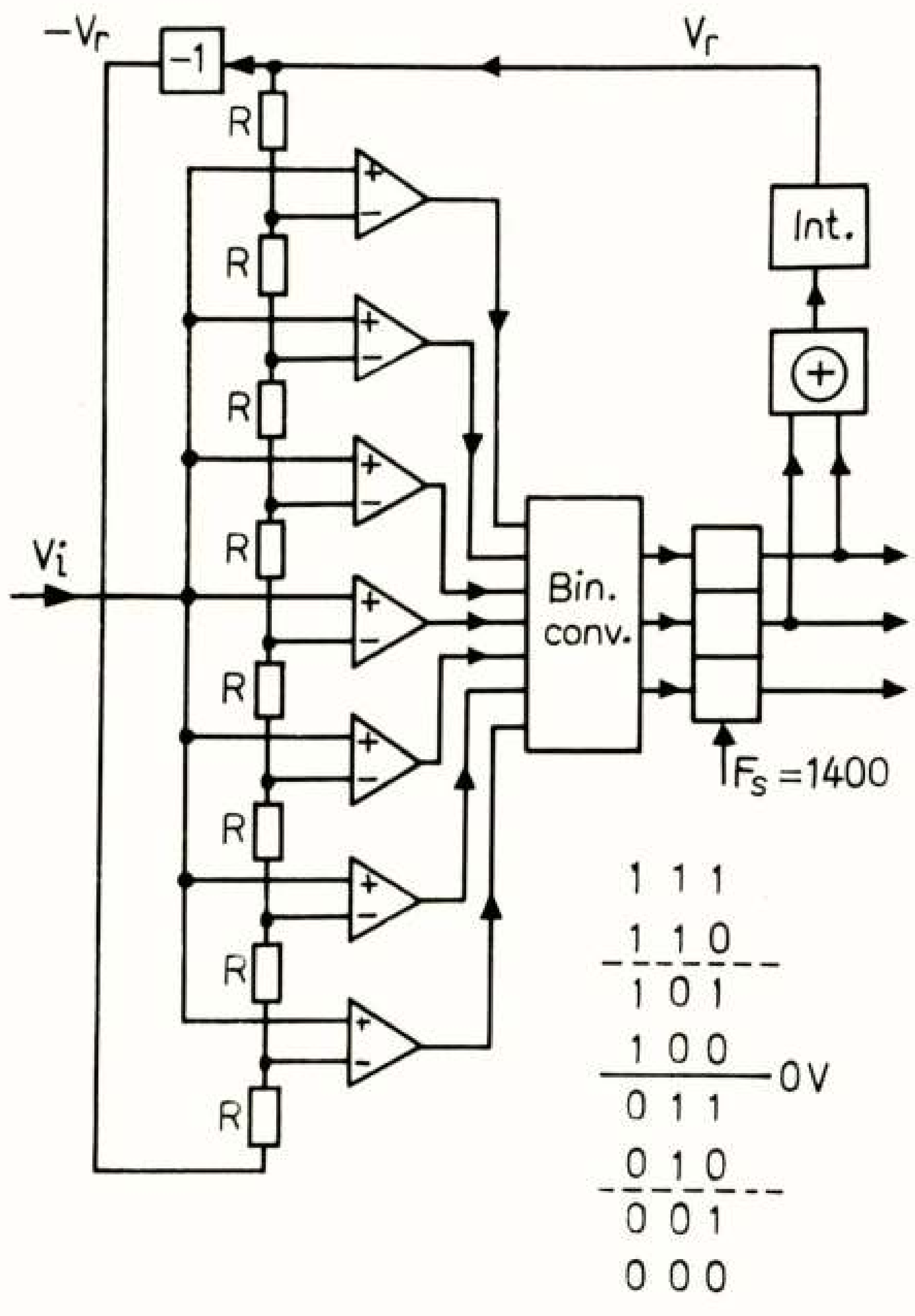


Fig. 6 Vereenvoudigd schema van de APCM-coder.

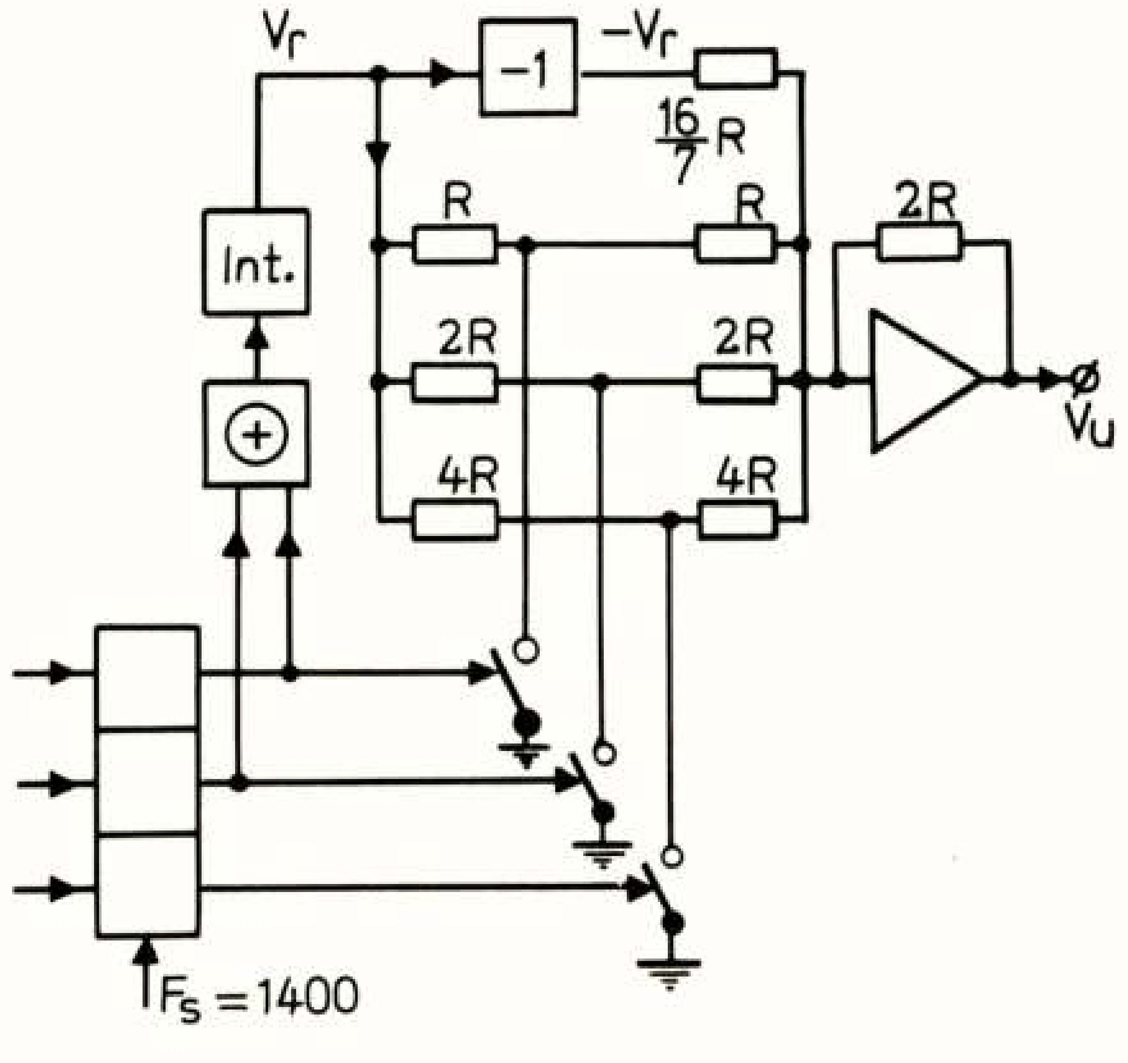


Fig. 7 Vereenvoudigd schema van de APCM-decoder.

bemonsteringstijdstippen getriggerd met een 1400 Hz kloksignaal F_s. De regelspanning V_r wordt opgewekt m.b.v. een exclusive - or schakeling en een integrator. De in Fig. 6 aangegeven tabel van de mogelijke 3-bits woorden bevat twee stippellijnen welke symbolisch drempels weergeven. Als er een APCM-woord wordt gecodeerd (in het uitgangsregister verschijnt) dat buiten de drempels ligt, dan geeft de exclusive-or schakeling een impuls af die aan de integrator wordt toegevoerd. Hierdoor neemt V_r toe. Wordt er een APCM-woord gecodeerd dat binnen de drempels valt dan wordt er geen impuls aan de integrator toegevoerd en V_r neemt hierdoor af. De tijdkonstante van de integrator is ong. 6 msec. en de tijdsduur van de impulsen die door de exclusive-or schakeling wordt afgegeven is 1/1400 sec. De regelspanning zal de omhullende van het ingangssignaal volgen, zodat ook de kwantiseringsniveau's in grootte meevariëren met de amplitude van het ingangssignaal.

De digitale multiplexer (zie Fig. 2) voegt de twee bitstromen van 4200 bps en 600 bps samen tot een synchrone bitstroom van 4800 bps. In de ontvanger worden deze bitstromen weer opgesplitst d.m.v. de demultiplexer. De synchronisatie van de demultiplexer laten we buiten beschouwing.

In de APCM-decoder, zie Fig. 7, wordt op dezelfde wijze als in de APCM-coder, de regelspanning V_r weer opgewekt. Het uitgangssignaal van de decoder wordt verkregen door V_r te vermenigvuldigen met de waarde van het

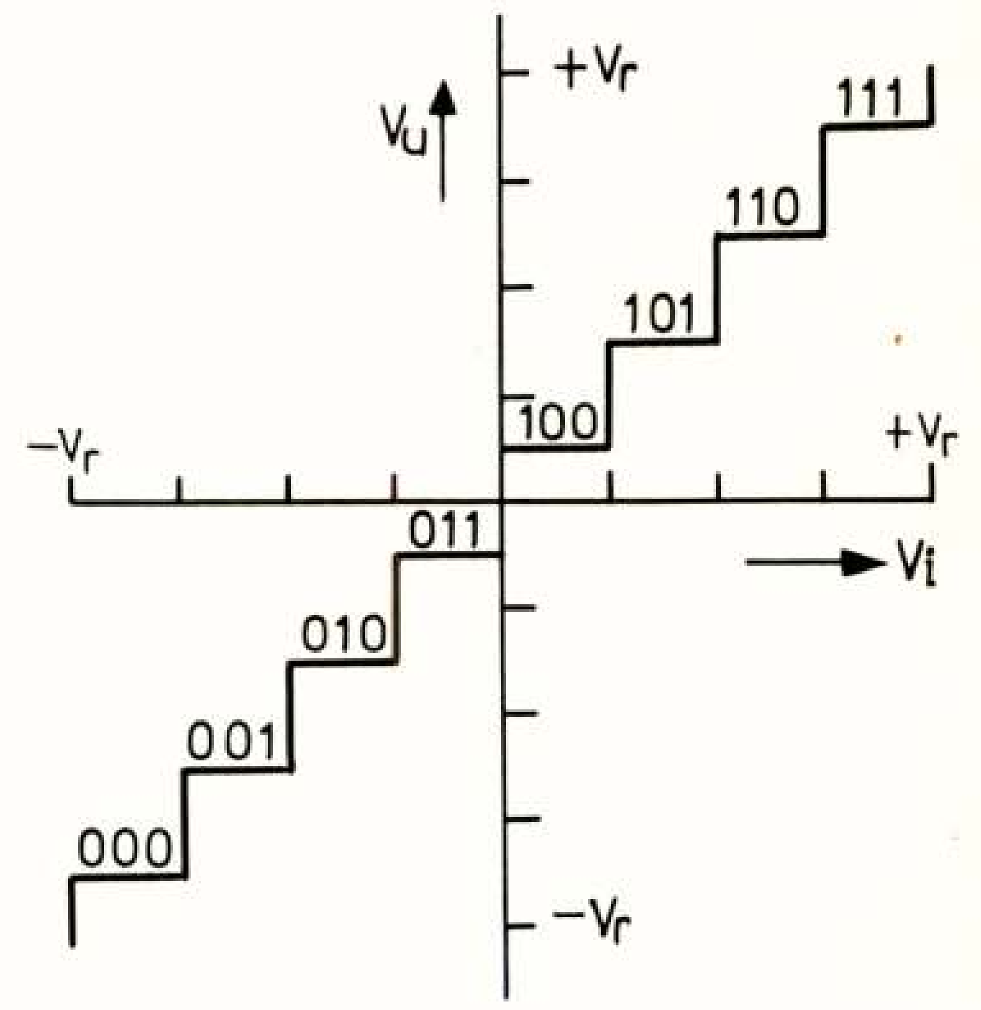


Fig. 8 Codeer-karakteristiek van het APCM-systeem.

ontvangen APCM-woord. Dit kan men eenvoudig realiseren met een weerstandsnetwerk en drie schakelaars, zoals in Fig. 7 is weergegeven. De verkregen codeer-karakteristiek van het APCM systeem is weergegeven in Fig. 8. V_i stelt het ingangssignaal van de coder voor en V_u het uitgangssignaal van de decoder.

De lineaire PCM-decoder genereert de betreffende regelspanningen. De analoge demultiplexer verdeelt deze regelspanningen naar de juiste kanalen.

RESULTATEN

Bij het beoordelen van de kwaliteit van de vocoder maken we onderscheid tussen de verstaanbaarheid, de natuurlijkheid of de herkenbaarheid van de spreker en de gevoeligheid van de vocoder voor transmissiefouten.

Uitgebreide verstaanbaarheidsproeven zijn niet gedaan, maar informele luisterproeven tonen aan dat de verstaanbaarheid ronduit goed is. Dit wordt vooral verkregen door het rechtstreeks overzenden van het eerste formantsignaal en het medesynthetiseren van de stemloze klanken. De aanwezigheid van het oorspronkelijke eerste formantsignaal aan de ontvangzijde draagt ook grotendeels bij tot de natuurlijkheid.

In feite zijn in deze vocoder twee klassieke vocoder problemen omzeild. Deze problemen hebben betrekking op het verkrijgen van het juiste excitatiesignaal voor de synthetisator. Enerzijds moet de synthetisator aangestoten worden met impulsen die in exact de goede grondtoonfrekwentie opgewekt moeten worden wanneer stemhebbende klanken gesproken worden. Anderzijds, bij stemloze klanken, zou men de synthetisator moeten aanstoten met ruis. Om dit te bewerkstelligen heeft men een detektor nodig die aangeeft of de gesproken klank stemloos of stemhebbend is. Bovendien heeft men voor het geval dat de klank stemhebbend is, een detektor nodig die aangeeft hoe groot de grondtoonfrekwentie precies is. Beide detectors zijn niet betrouwbaar te realiseren zonder zeer ingewikkelde apparatuur. In de beschreven vocoder worden de verschillende excitatiesignalen niet op verschillende tijdstippen aan de synthetisator toegevoerd, maar gelijktijdig in verschillende frekwentiebanden. Wanneer een klinker wordt ingesproken, wordt aan de ontvangkant in het derde formantkanaal ruis gegenereerd op een laag niveau, overeenkomstig de sterkte van de relatief zwakke derde formant. De signalen in de eerste en tweede formantkanalen, welke beide de goede grondtoon bevatten, domineren nu. Bij stemloze klanken daarentegen domineert het ruissignaal uit het derde formantkanaal. Op deze wijze verkrijgt men een "gladde" overgang van stemhebbende naar stemloze klanken en omgekeerd.

De gevoeligheid voor transmissiefouten is gering. Bij een foutenkans van 10^{-3} in de overgezonden bitstroom merkt men nauwelijks enige invloed. Bij een foutenkans van 3.10^{-3} merkt men verstoringen op, speciaal tijdens de spraakpauzes. Bij een foutenkans van 10^{-2} is het systeem nog steeds bruikbaar, maar storing in het spraaksignaal is duidelijk aanwezig. Bij foutenkansen van 3.10^{-2} en hoger, wordt het spraaksignaal zodanig aangetast dat de verstaanbaarheid grotendeels verloren gaat.

LITERATUUR

1) Flanagan, J.L. Speech Analysis Synthesis and Perception. Springer-Verlag, Berlin (1965).
2) Flanagan, J.L. Resonance-vocoder and Baseband Complement. IRE Trans. on Audio AU-8, pp. 95-102(1960).
3) Peterson, G.E. and Barney, H.L. Control Methods used in a Study of the Vowels. J. Acoust. Soc. Am. 24, pp. 175-184 (1952).
4) Wilkinson, R.M. An Adaptive Pulse Code Modulator for Speech. IEEE Int. Conf. on Comm. Montreal, pp.1-11 1-15 (1971).

Voordracht gehouden op 12 mei 1976 in het Instituut voor Zintuigfysiologie TNO te Soesterberg op een gemeenschappelijke vergadering van het NERG (no. 256), de Benelux-section IEEE en het Nederlands Akoestisch Genootschap.

REDUCTIE VAN SPRAAKBANDBREEDTE MET BEHULP VAN EEN VOCODER DIE GEBRUIK MAAKT VAN VOORUITREGELING

Ir. K.E. Kuijk, Dr.Ir.N.V. Franssen
Philips Research Laboratories
Eindhoven,The Netherlands

A new type of a voice-excited vocoder is described. A part of the receiving system has been copied in the transmitter system, so that signals can be generated for control of artificial frequency bands in order to agree in energy level with the corresponding frequency bands of the original speech.

Inleiding

Voor een goede verstaanbaar van spraak is tenminste een bandbreedte van 300 - 3200 Hz noodzakelijk. In vele gevallen, b.v. bij transatlantische telefonie, zou men deze bandbreedte nog willen reduceren.

Dit kan met een vocoder. In dit artikel wordt een zogenaamde voice-excited vocoder beschreven. In een dergelijk systeem wordt een deel van de spraakband rechtstreeks overgezonden. Aan de ontvangzijde wordt dit deel dan tevens gebruikt om het overige deel van het spectrum te reconstrueren.

1. De opbouw van spraakklanken

Spraak bestaat uit een reeks elkaar opvolgende klanken; klinkers en medeklinkers.

Deze klanken kunnen een grote mate van periodiciteit hebben - ze hebben dan een duidelijke toonhoogte - of niet, b.v. de letters s,t,f,p,g,k of gefluisterde spraak.
In beide gevallen echter blijkt de trillingsenergie zich hoofdzakelijk in een tweetal frekwentie domeinen te bevinden, de zogenaamde formanten, die voor een bepaalde spraakklank karakteristiek zijn. Zo geeft fig. 1 de ligging aan van de formanten van de klinkers die elkaar gedeeltelijk overlappen.

Daar wij onze mondstand n.l. slechts betrekkelijk langzaam kunnen veranderen, zal ook de spectrale inhoud van spraak slechts langzaam veranderen. Daar bovendien ons gehoororgaan over tertsbandbreedtes integreert zal exacte overdracht van de frequenties niet nodig zijn maar is reproductie van formanten binnen de nauwkeurigheid van een tertsbandbreedte voldoende om verstaanbare spraak te verkrijgen. Hierin ligt nu een mogelijkheid voor bandbreedte reductie.

Wel is vaak het onderscheid "toonhoogte hebbend of niet" moeilijk automatisch vast te stellen. De moeilijkheden hiervan zijn te omzeilen door een voldoend groot gedeelte van de spraakband rechtstreeks over te zenden en hieruit met behulp van aan de zendzijde gevormde regelsignalen het gehele spraakspectrum te reconstrueren, zodat de periodiciteit behouden blijft. Dit noemt men een voice-excited vocoder.

2. Het principe van de spraakband reductie

Het hier te beschrijven systeem is zo een voice-excited vocoder. Als direct over te zenden band nemen we het gebied van 200 - 800 Hz, waarvan de onderste grens gegeven wordt door toevallig aanwezige filters; hierin is dan praktisch altijd de eerste formant aanwezig.

Zowel aan de zend- als ontvangstzijde wordt deze grondband op onderling identieke wijze vervormt teneinde hieruit een frequentie band van 800 - 3200 Hz af te leiden die zo goed mogelijk gelijkmatig gevuld is. Dit signaal heeft dan dezelfde periodieke structuur als de oorspronkelijke spraak daar het uit het belangrijkste gedeelte hiervan is afgeleid. Dit lukt heel aardig door de 200 - 800 Hz band eerst dubbel gelijk te richten; via een capacitieve koppeling peak-clipping toe te passen en vervolgens dit signaal te differentiëren.

Aan de zendzijde worden nu met behulp van tweemaal een zestal tertsfilters zowel het spraakgedeelte van 800 - 3200 Hz als het kunstmatige signaal in telkens zes banden opgesplitst. De gemiddelde amplitude aan de uitgangen van overeenkomstige tertsfilters worden nu onderling vergeleken en hieruit worden regelsignalen afgeleid. Ook aan de ontvanstzijde wordt het kunstmatige signaal met behulp van tertsfilters in de zes overeenkomstige tertsbanden gesplitst. Daar de kunstmatige band aan de ontvangstzijde door de onderling gelijke bewerking exact gelijk is aan die

van de zendzijde, zijn de regelsignalen een nauwkeurige aanwijzing voor wat er aan de ontvangstzijde aan de gemiddelde amplitudes van de signalen uit de tertsbanden moet veranderen om dezelfde amplitudes te verkrijgen als van de signalen uit de tertsbanden van de oorspronkelijke spraak.

Daar de amplitudes in de tertsbanden niet snel kunnen veranderen kunnen de regelsignalen begrensd worden tot een bandbreedte van minder dan 50 Hz. Men kan dus volstaan met het overzenden van de oorspronkelijke grondband van 200 - 800 Hz en de zes regelsignalen. Aan de ontvangstzijde worden de gemiddelde amplitudes uit de zes tertsbanden met behulp van regelversterkers door de corresponderende regelsignalen geregeld. Na samenvoeging van de grondband en de zes aldus gereconstrueerde deelbanden heeft men weer een zeer verstaanbare spraak herkregen die met slechts ongeveer 1000 Hz bandbreedte is overgezonden.

3. Het amplitude regel systeem

Tot de belangrijkste onderdelen van het systeem van Fig. 2 behoren de regelcircuits voor de amplitude.

Het eerste voorstel over dit uit te voeren is aangegeven in Fig. 3 en maakt gebruik van een vermenigvuldiger in een terugkoppellus. Het signaal V_A wordt gelijkgericht en na filtering in een laagdoorlaatfilter wordt een spanning evenredig met de amplitude $\hat{V}_A$ (voor sinusvormige signalen) verkregen. Het signaal V_B wordt via een vermenigvuldiger in een overeenkomstige schakeling gestuurd. De beide signalen overeenkomend met de amplitudes, worden vergeleken in een verschilversterker en deze stuurt de vermenigvuldiger tot er geldt $M \times \hat{V}_B = \hat{V}_A$. Het nadeel van deze methode is dat de versterking in de regellus verandert als funktie van de amplitudes. Daardoor is de nauwkeurigheid niet konstant, en tevens kan de maximale open lus versterking niet te groot zijn wegens gevaar voor parasitair oscilleren. Om de rimpel op het regelsignaal V_r klein te houden zijn er tweede of hogere orde laagdoorlaatfilters nodig en deze geven ook aanleiding tot oscilleren van de regellus; dit kan wel gedeeltelijk voorkomen worden door dubbelzijdige gelijkrichting toe te passen.

Een veel direktere methode, die al deze nadelen niet heeft, is aangegeven in Fig. 4. Hierbij worden beide signalen gelijkgericht in een (enkelzijdige) gelijkrichter en via laagdoorlaatfilters van tweede- of hogere orde worden de amplitude informaties in de vorm van gelijkspanningen verkregen. Veronderstellen we voor het gemak dat $V_A = \hat{V}_A \cos\omega_A t$ en $V_B = \hat{V}_B \cos\omega_B t$, dan zijn deze gelijkspanningen dus evenredig met $\hat{V}_A$ en $\hat{V}_B$. Vervolgens worden hiervan in logarithmische omzetters spanningen gemaakt evenredig met $\ln \hat{V}_A/V_C$ en $\ln \hat{V}_B/V_C$, waarbij V_C een referentiespanning is, en deze worden van elkaar afgetrokken, zodat er een spanning evenredig met $\ln \hat{V}_A/\hat{V}_B$ ontstaat. Als hiermee in een exponentiële vermenigvuldiger het signaal $V_B = \hat{V}_B \cos\omega_B t$ gestuurd wordt, is het resultaat $\hat{V}_B \cos\omega_B t . e \exp \ln \hat{V}_A/\hat{V}_B = \hat{V}_A \cos\omega_B t$, d.w.z. het signaal met frekwentie ω_B krijgt de amplitude $\hat{V}_A$ van het signaal met frekwentie ω_A.

In de experimentele schakeling, opgezet om de verstaanbaarheid te toetsen, is het signaal $\ln \hat{V}_A/\hat{V}_B$ als regelsignaal gebruikt. Bij b.v. een telefoniesysteem is het niet goed mogelijk om dit signaal $\ln \hat{V}_A/\hat{V}_B$ over te zenden, omdat een eventuele verzwakking op de lijn aanleiding geeft tot een signaal aan de uitgang van de exponentiële vermenigvuldiger (aan de ontvangkant):

$$\hat{V}_B \cos\omega_B t . \; e^{\alpha \hat{V}_A/\hat{V}_B} = \left(\frac{\hat{V}_A}{\hat{V}_B}\right)^{\alpha} . \; \hat{V}_B \cos\omega_B t$$

en dus niet tot het beoogde resultaat leidt. Deze moeilijkheid is te vermijden door een pilot signaal mee over te zenden die een bepaalde bekende amplitude moet hebben aan de ontvangstzijde en dus voor bijregeling kan zorgen of door het signaal $\ln \hat{V}_A/\hat{V}_B$ eerst op een exponentiële versterker te zetten. Het signaal evenredig met $\hat{V}_A/\hat{V}_B$ dat dan ontstaat, kan worden overgezonden. In de ontvanger wordt het signaal $\alpha \hat{V}_A/\hat{V}_B$ op een logarithmische omzetter gezet; het uitgangssignaal $\ln \alpha V_A/V_B$ wordt gebruikt als regelsignaal in een exponentiële vermenigvuldiger om het signaal $\hat{V}_B \cos\omega_B t$ te sturen. Er ontstaat dan $\alpha \hat{V}_A \cos\omega_B t$, d.w.z. het gewenste resultaat, afgezien van een zekere verzwakking α.

In het volgende gedeelte zal de manier, waarop de elektronische deelschakelingen gerealiseerd zijn, besproken worden.

4. De gelijkrichtschakeling

In Fig. 5 is de gebruikte gelijkrichtschakeling getekend. Het is een enkelzijdige gelijkrichtschakeling. De schakeling is in de praktijd belast met een 2e orde laagdoorlaatfilter, waarvan hier de eerste sektie is getekend. Om nu het opladen en ontladen van de kondensator via eenzelfde weerstand te doen verlopen is een emittervolger i.p.v. een diode opgenomen in de terugkoppelleiding. De nauwkeurigheid van deze gelijkrichter is ongeveer 5% over een gebied van 73 dB (1 mV - 5 V).

5. De logarithmische omzetters

In Fig. 6 is het schema van een logarithmische omzetter getekend. Voor de transistoren Q_1 en Q_2 geldt de exponentiële transistor karakteristiek

$$I_{c1} = I_{o1}\cdot(e^{qV_{be1}/k.T_1}-1) \cong I_{o1}\cdot e^{qV_{be1}/k.T_1}$$

$$I_{c2} = I_{o2}\cdot(e^{qV_{be2}/k.T_2}-1) \cong I_{o2}\cdot e^{qV_{be2}/k.T_2} \qquad (1)$$

met : k = konstante van Boltzmann

T = abs. temp.

q = lading van het electron.

en waarbij $I_{01.2}$ de verzadigingsstroom is van resp. Q_1 en Q_2. Als de ingangsstroom I_2 van de operationele versterker verwaarloosbaar is, geldt er :

$$I_{c1} = i_1 = \frac{\hat{V}_A}{R_4} \qquad (2)$$

zodat

$$V_{be1} = \frac{k.T_1}{q}\ \ln \frac{\hat{V}_A}{I_{o1}.R_4} \qquad (3)$$

Een analoge formule kunnen we afleiden voor de bewerking van het signaal V_B :

$$V_{be2} = \frac{k.T_2}{q}\ \ln \frac{\hat{V}_B}{I_{o2}.R_4'} \qquad (4)$$

Als nu de transistoren Q_1 en Q_2 een paar identieke geintegreerde transistoren zijn, vlak bij elkaar in één I.C., dan geldt dat I_{o2} en I_{o2} goed aan elkaar gelijk zijn, met een onnauwkeurigheid van 4%, en dat de temperaturen T_1 en T_2 aan elkaar gelijk zijn. In dat geval geldt, als $R_4 = R_4'$:

$$V_s = V_{be1} - V_{be2} = \frac{k.T_1}{q}\ \ln \frac{\hat{V}_A}{\hat{V}_B} \qquad (5)$$

In de experimentele schakeling is dit signaal V_s als regelsignaal voor de experimentele vermenigvuldiger gebruikt.

6. De exponentiële vermenigvuldiger

De exponentiële vermenigvuldiger is aangegeven in Fig. 7. In deze schakeling is Q_3, Q_4 weer een geïntegreerd transistorpaar. Er geldt weer de exponentiële transistorkarakteristiek :

$$I_{c3} \simeq I_{o3}\cdot e^{qV_{be3}/k.T_3}$$

$$I_{c4} \simeq I_{o4}\cdot e^{qV_{be4}/k.T_4} \qquad (6)$$

waarbij weer geldt

$$I_{o3} = I_{o4}\ ,\quad T_3 = T_4$$

dus

$$I_{c4} = I_{c3}\cdot e^{q(V_{be4}-V_{be3})/k.T_3} = I_{c3}\cdot e^{qV_r/k.T_3} \qquad (7)$$

In de experimentele situatie is het regelsignaal V_R gelijk aan V_S uit (5) : $V_R = V_S$. Tevens is in dat geval $T_3 = T_1$, dus geeft (5) in (7) :

$$I_{c4} = I_{c3}\cdot \frac{\hat{V}_A}{\hat{V}_B} \qquad (8)$$

Is $I_{c3} = I_3 + \frac{V_B}{R_6}$ (9) waarbij V_B het wisselspanningssignaal is, dan is het kollektorsignaal V_{c4}

$$V_{c4} = R_{c4}\cdot I_{c4} = R_{c4}\cdot(I_3 + \frac{V_B}{R_6})\cdot \hat{V}_A/\hat{V}_B \qquad (10)$$

Het wisselspanningssignaal dat door de CR-koppeling wordt doorgelaten is :

$$V_u = \frac{R_{c4}}{R_6}\cdot V_B\cdot \hat{V}_A/\hat{V}_B \qquad (11)$$

d.w.z. het signaal V_B krijgt de amplitude van het signaal V_A, zoals de opzet was. In deze vorm is de vermenigvuldiger dus in de experimentele situatie uitgevoerd. (Fig. 8). In een praktische situatie d.w.z. met tussenschakeling van een telefonielijn, ontbreekt in de vermenigvuldiger aan de zendzijde weerstand R_6. Dan wordt vergelijking (10) :

$$V_{c4} = R_{c4}\cdot I_3\cdot \hat{V}_A/\hat{V}_B \qquad (12)$$

en dit signaal kan wel overgezonden worden.
Aan de ontvangzijde wordt $\propto V_{c4}$ (met een eventuele verzwakking $\propto$) weer op een logarithmische omzetter gezet van de vorm zoals besproken in Hoofdstuk V, tezamen met een hulpspanning $R_{c4}'I'_3$ ($= R_{c4}I_3$). Er ontstaat dan een spanningsverschil aan de uitgang van de logarithmische omzetter gelijk aan $\ln \propto V_A/V_B$ welke weer gebruikt wordt als regelsignaal om een exponentiële vermenigvuldiger gelijk aan die uit Fig. 8 te sturen.

7. Meetresultaten

In Fig. 9 is het gebied aangegeven waarin de twee signalen V_A en V_B zich mogen bewegen, zoals gemeten in de experimentele situatie, d.w.z. met de logarithmische omzetters direkt aangesloten op de exponentiële vermenigvuldigers (Fig. 9).

Om de onderlinge gelijkheid van de regelsystemen na te gaan, werd een in frequentie variërende sinusvormige trilling gelijktijdig en in gelijke grootte aan de drie tertsfiltersets toegevoerd. Fig. 10 geeft de responsie aan de uitgang van de ontvangstzijde aan.

Tenslotte kan met een afluistertest van een onbekend fragment spraak de verstaanbaarheid van het gereconstrueerde spraaksignaal nagegaan worden.

8. "Acknowledgment"

De auteurs zouden graag de heren J. Hansma

en J.W. Siebelink bedanken voor hun onderzoek
aan de regelsystemen en de heer M.H. Geelen voor de
bouw van en de metingen aan het totale systeem.

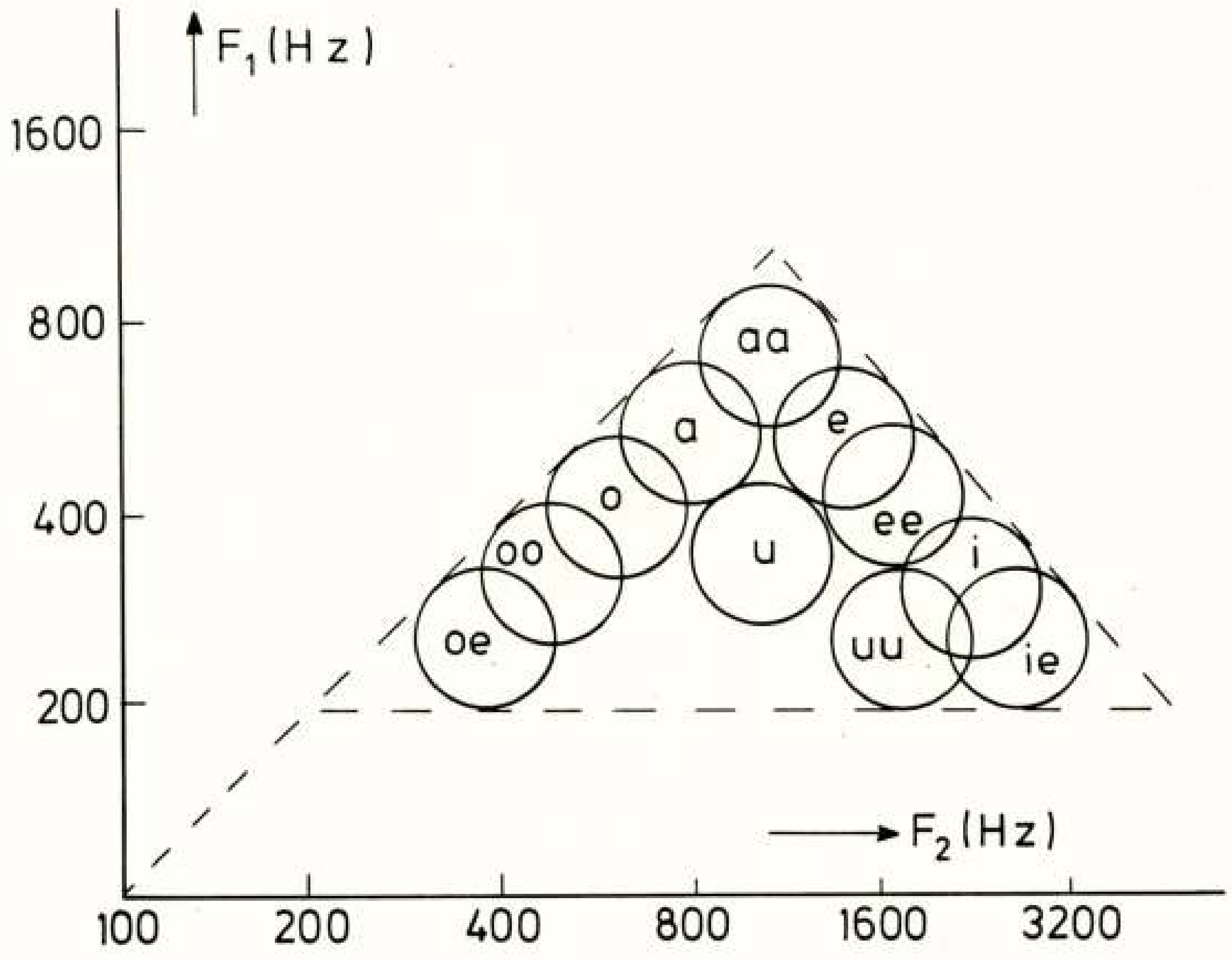


Fig.1 FREQUENTIE GEBIEDEN VAN KLINKERS

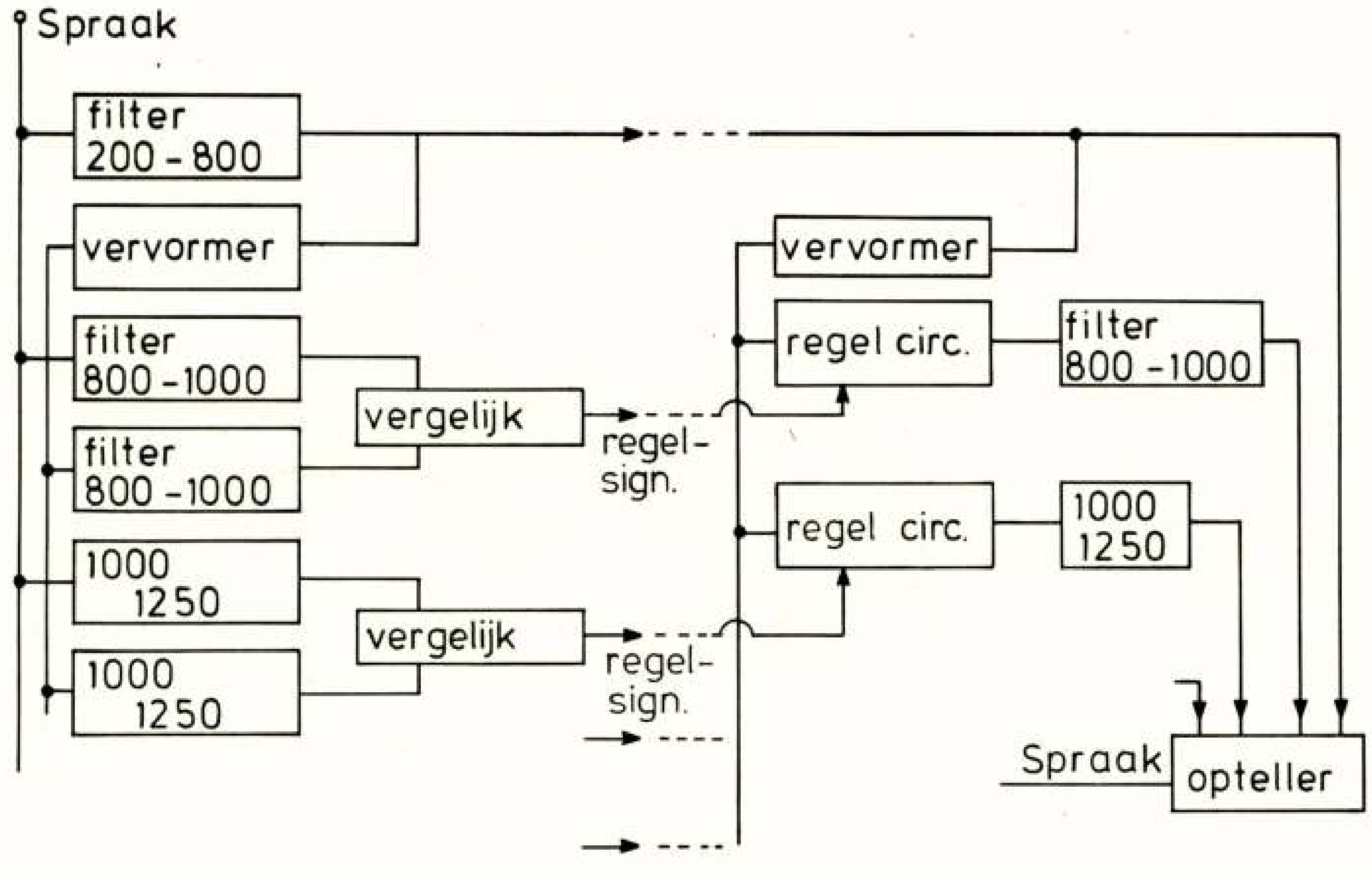


Fig.2 PRINCIPE SCHEMA

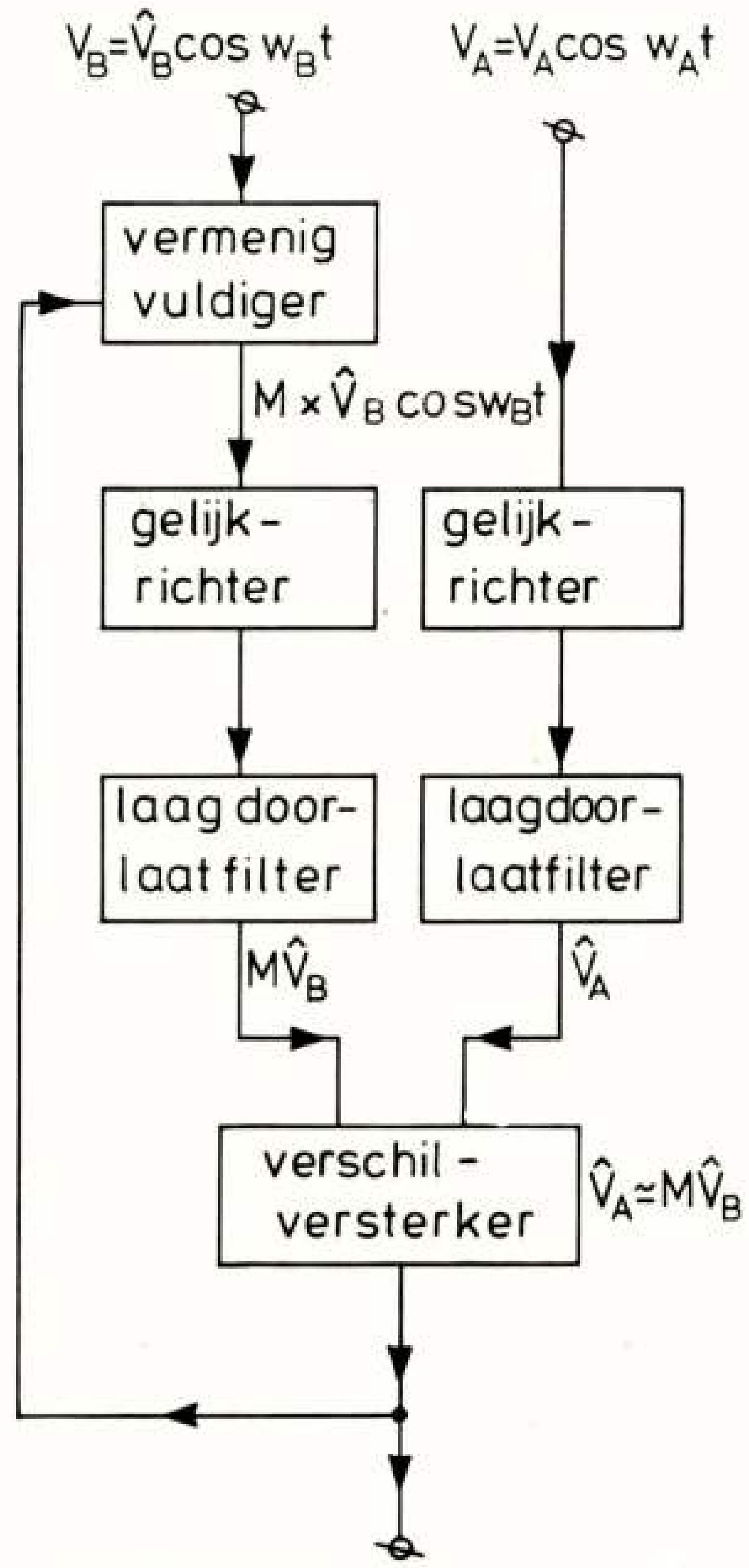


Fig. 3 REGELSYSTEEM MET TERUG KOPPELING

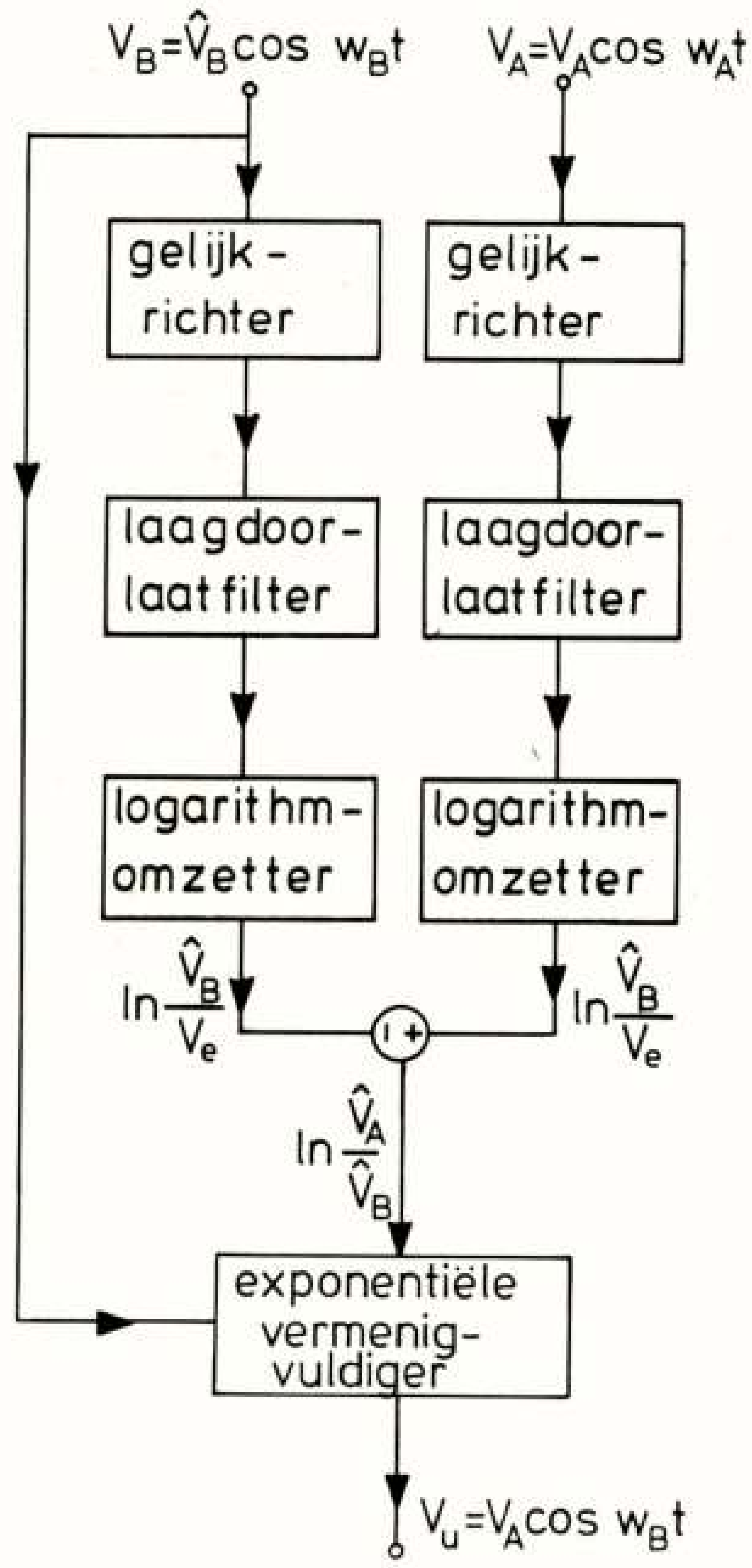


Fig. 4 VOORUIT REGELING VAN AMPLITUDE

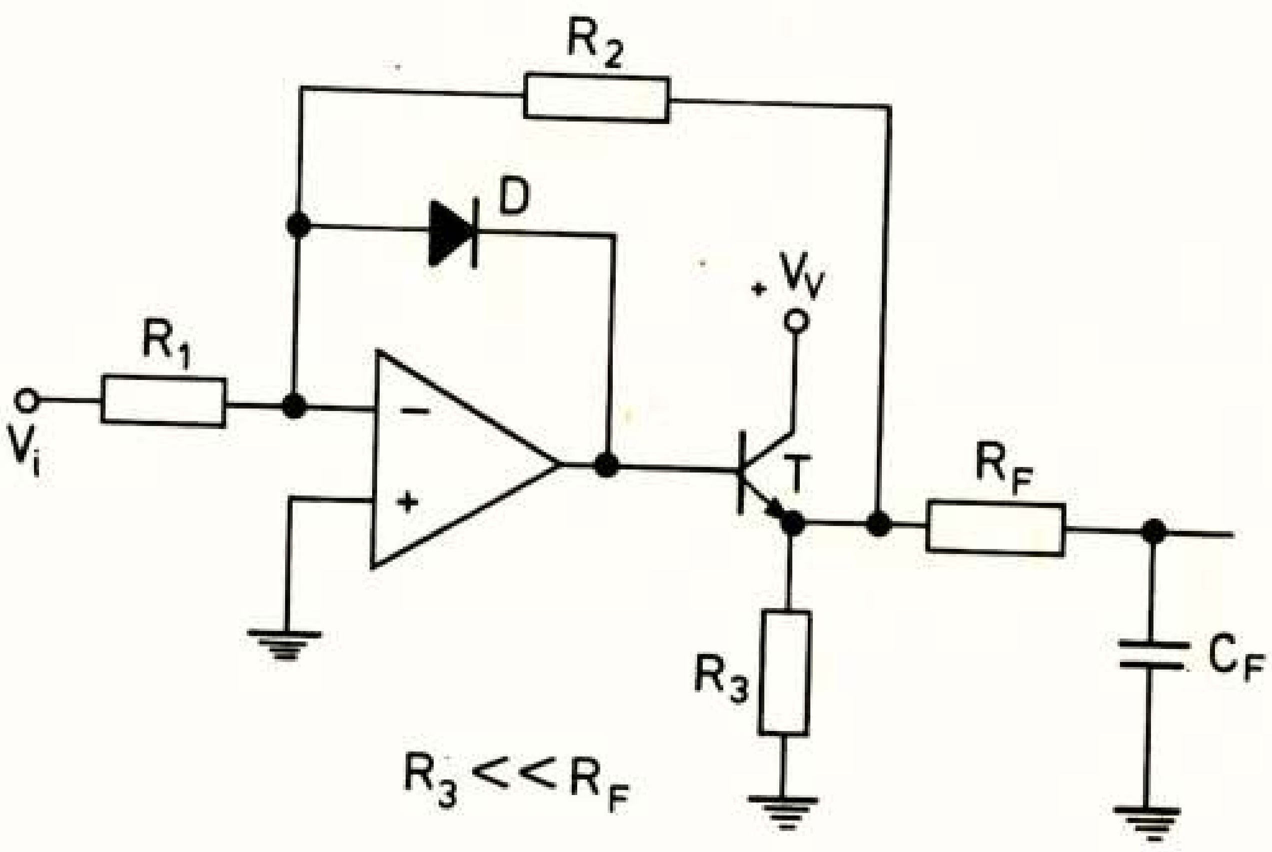


Fig. 5 DE GELIJK RICHTERSCHAKELING

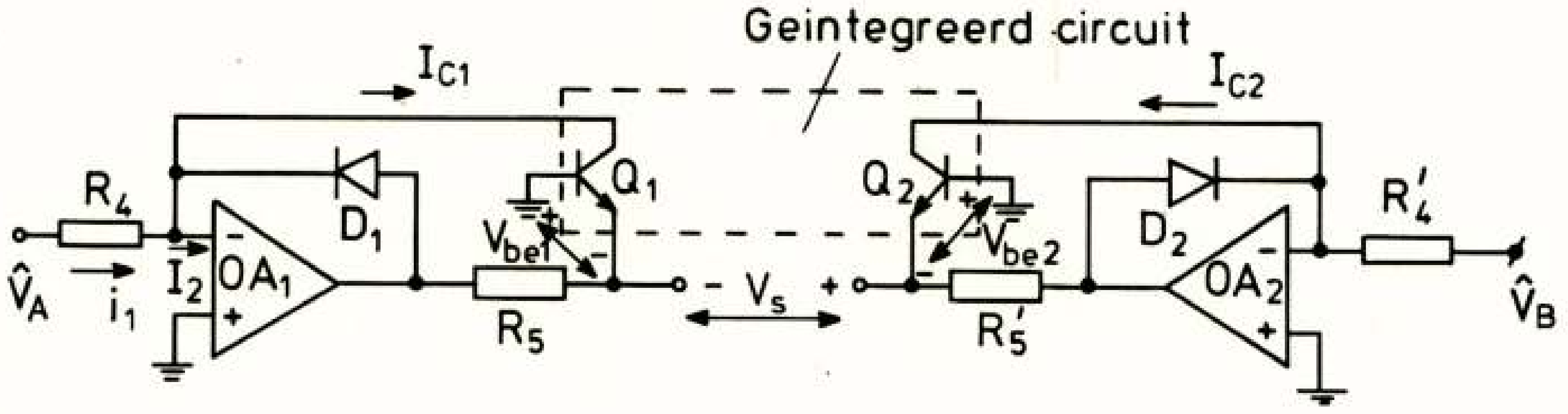


Fig. 6 DE LOGARITHMISCHE OMZETTERS

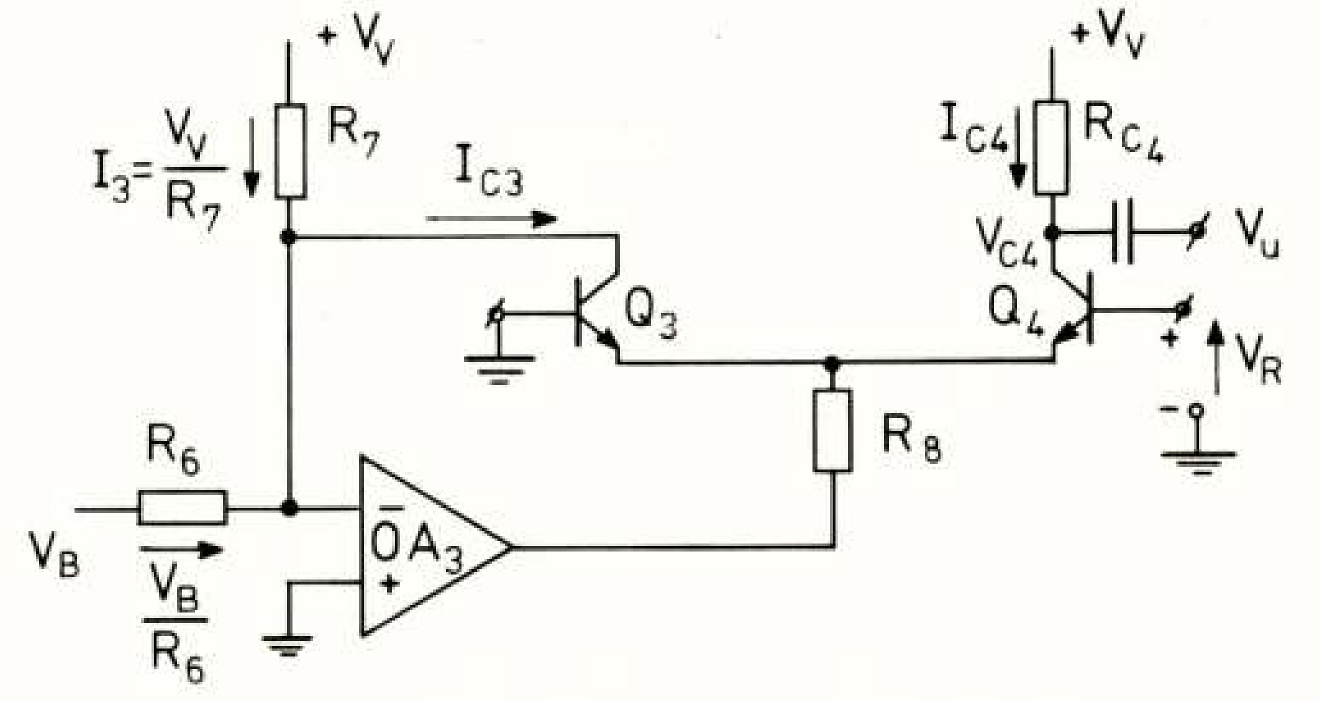


Fig. 7 DE EXPONENTIËLE VERMENIGVULDIGER

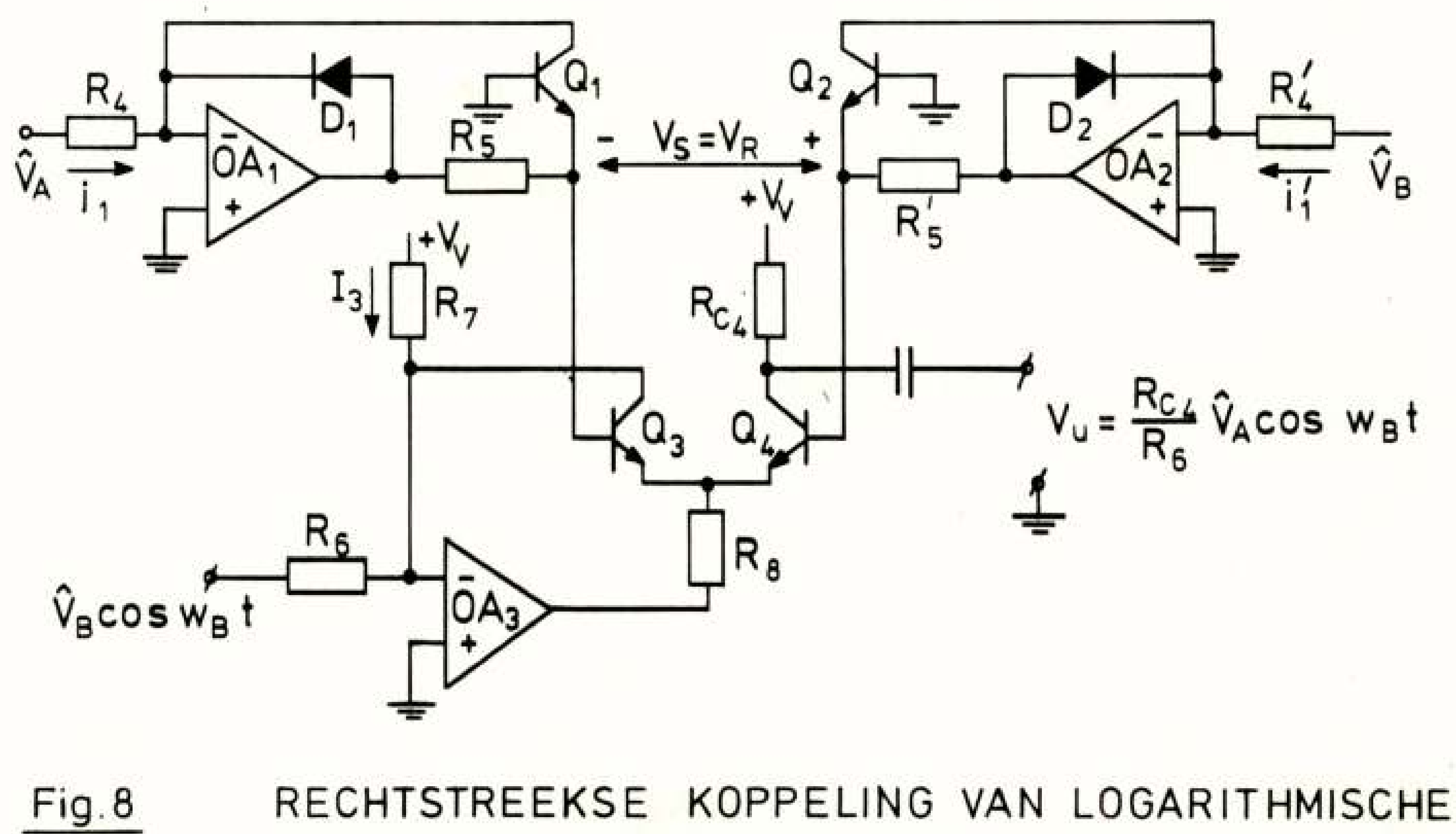


Fig. 8 RECHTSTREEKSE KOPPELING VAN LOGARITHMISCHE OMZETTERS EN EXPONENTIELE VERMENIGVULDIGER

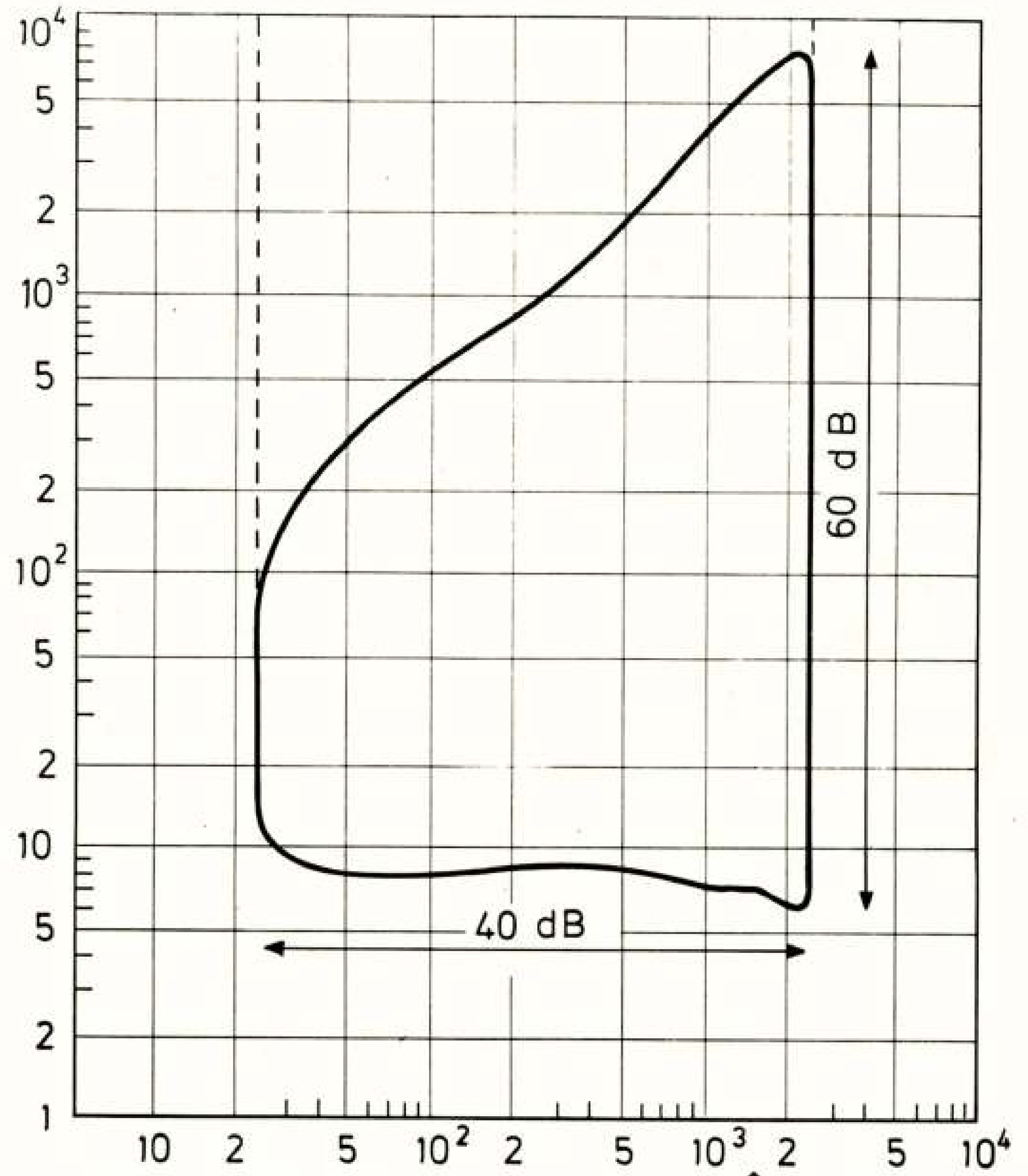


Fig. 9 Regel gebied → $\hat{V}_B$ (mV)

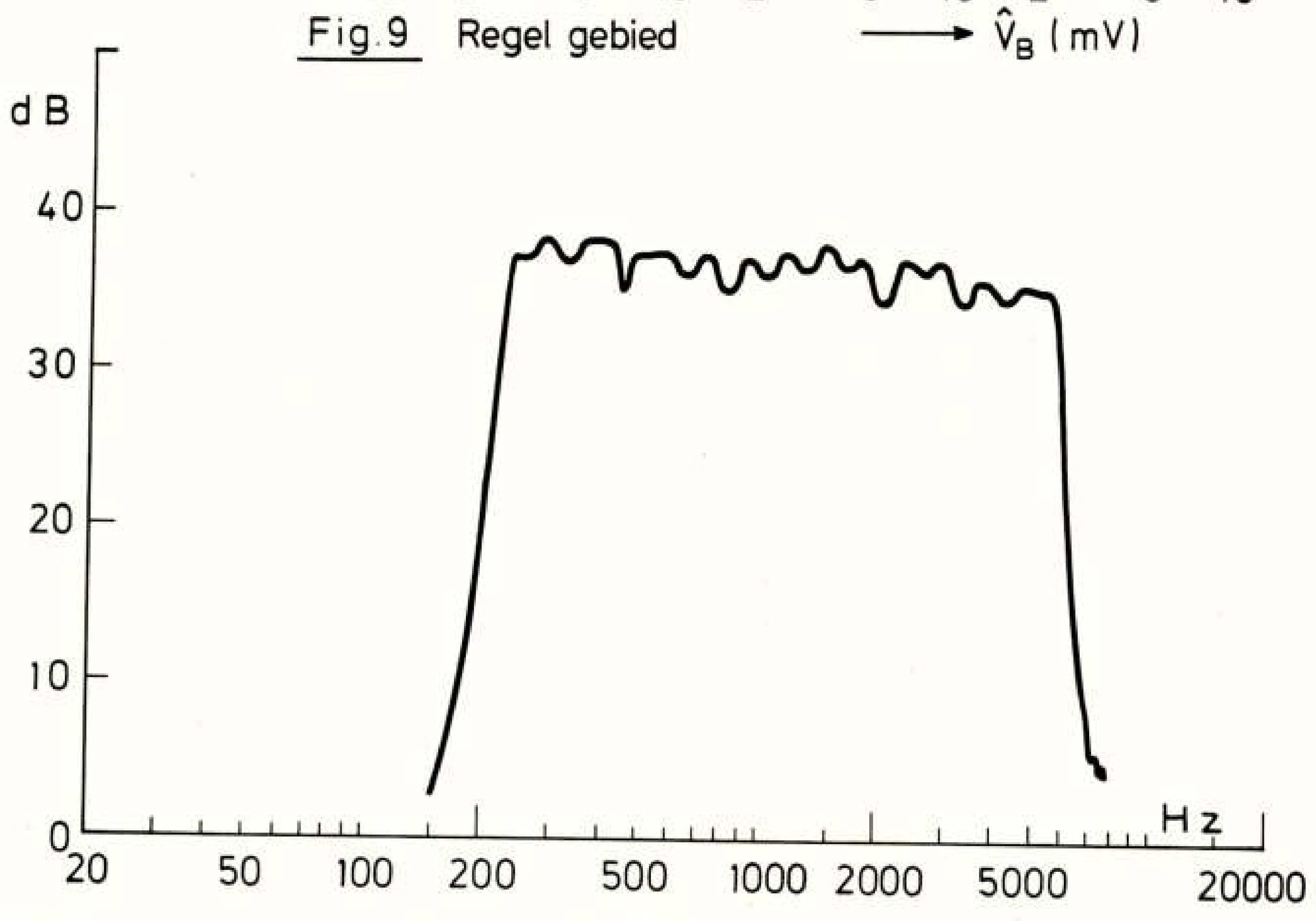


Fig. 10 GELIJKHEID VAN DE REGEL CIRCUITS

Voordracht gehouden op 12 mei 1976 in het Instituut voor Zintuigfysiologie TNO te Soesterberg op een gemeenschappelijke vergadering van het NERG (no. 256), de Benelux-section IEEE en het Nederlands Akoestisch Genootschap.

ADAPTIEVE COMPENSATIE VAN ECHO OP TELEFOONLIJNEN

Ir. H. van Noort

Dr. Neher Laboratorium

This paper deals with problems of echo and delay time as they are experienced on long international connections. The difficulties which arise with the present voice switched echo suppressors are shortly described. The problems caused by the voice-operated switches can be avoided with the adaptive echo-canceller, generating a signal equivalent to the echo and using this to cancel the real echo. The principles of the adaptive echo-canceller are explained and some information about a practical realisation is given.

I. ECHO OP TELEFOONVERBINDINGEN

In fig. 1 is een telefoonverbinding schematisch weergegeven. Om een lange afstand zonder al te veel demping te overbruggen wordt gebruik gemaakt van een versterkte verbinding (D-E). In de uiteinden van de verbinding worden om economische redenen onversterkte tweedraadsverbindingen toegepast (CD en EF). In

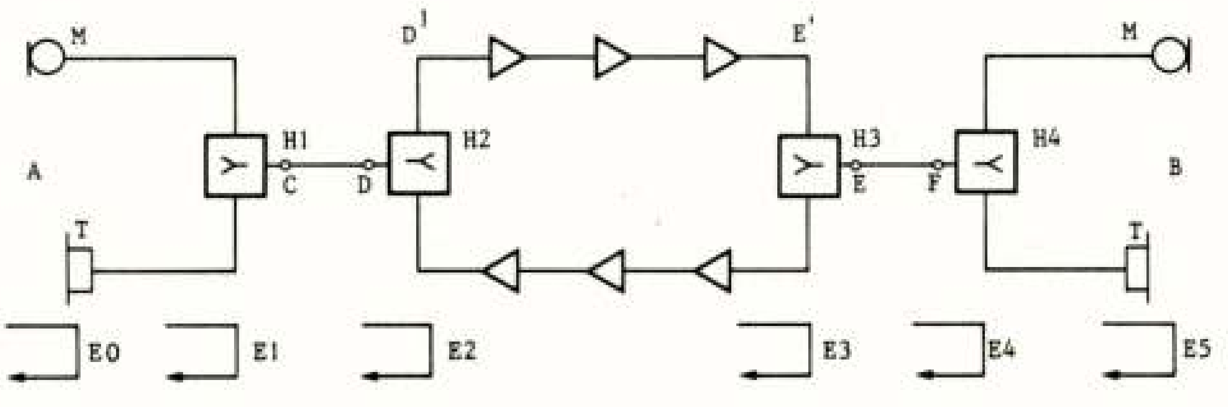


Fig. 1: Schematische weergave van een telefoonverbinding met echo's.

het toestel wordt deze tweedraadsverbinding weer vierdraads voor de microfoon (M) en de telefoon (T). De verschillende overgangen tussen vierdraads- en tweedraadsverbindingen worden gemaakt met de vorken $H_1 \dots H_4$. Dit zijn brugschakelingen waarbij de overloop van ene naar de andere vierdraadszijde meer gedempt wordt naarmate de impedantie van de tweedraadszijde meer gelijk is aan de impedantie van de vorkbalans.

De vorkbalans is een eenvoudig netwerkje, dat de gemiddelde impedantie van de tweedraadszijde benadert. De impedantie van de tweedraadszijde wordt o.a. gevormd door allerlei verschillende soorten kabels van verschillende lengten. Daardoor is er aanzienlijk verschil tussen de impedantie van de balans en van de tweedraadszijde met als gevolg dat signalen bij de vork overlopen die als echosignalen bij de spreker terugkomen. Wanneer iemand via de telefoon spreekt, hoort hij zijn eigen stem via de normale weg van mond tot oor (Eo) en via de reflecties E1 tot E5. Deze reflecties zijn niet hinderlijk zolang ze een verwaarloosbare vertraging hebben; dit is het geval voor Eo en E1 en meestal voor E2. De reflecties E3, E4 en E5 hebben echter een vertraging van meer dan twee maal de propagatietijd tussen D^1 en E^1 en deze tijd is niet altijd verwaarloosbaar. De hinderlijkheid van de echo is groter naarmate de vertragingstijd groter en de demping kleiner is. Een illustratie hiervan wordt gegeven in fig. 2, (Richards '63).

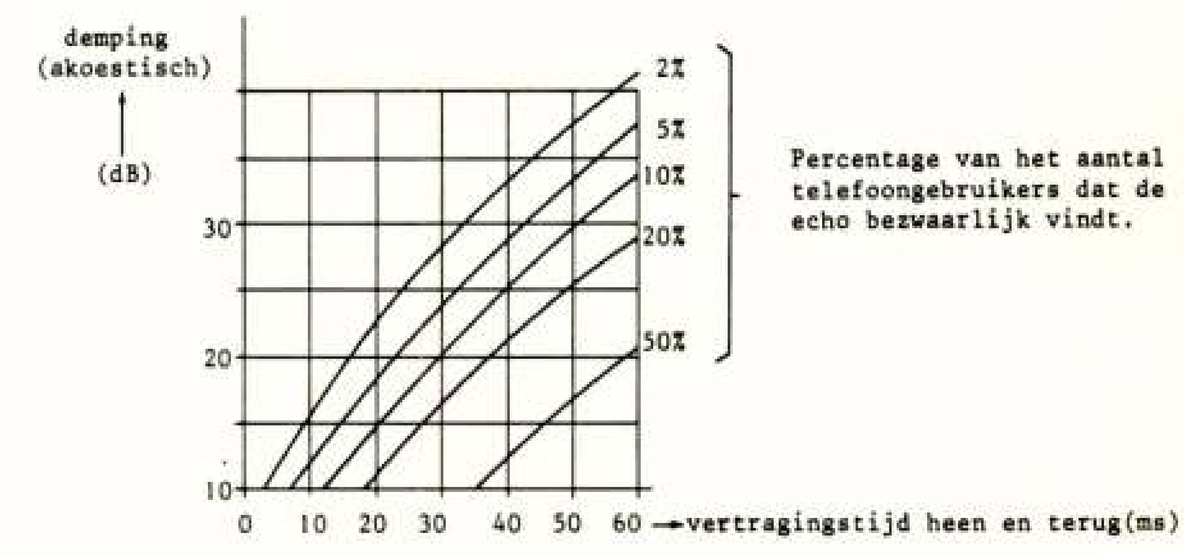


Fig. 2: Kwaliteitswaardering van verbindingen met echo.

In de praktijk bedraagt de overloopdemping van de vork 10 - 20 dB, voor de akoestische demping komt hier nog ± 4 dB bij voor het toestel en 2 - 14 dB voor de rest van de verbinding dus totaal 16 - 38 dB. Met deze gegevens blijkt uit fig. 2 dat maatregelen tegen echo nodig zijn voor vertragingstijden > 25 msec.

II. LOOPTIJD VAN TELEFOONVERBINDINGEN

Bij kabelverbindingen kunnen afstanden tot 20.000 km voorkomen, de propagatiesnelheid is ongeveer 165 km/sec zodat een looptijd van 120 msec (240 msec voor heen en terug) kan optreden. Voor een afstand Amsterdam - New York van ± 5000 km is de looptijd 35 msec (70 msec heen en terug). Voor de

telefoongebruiker is de looptijd heen en terug van belang zodat we in het vervolg alleen deze waarde zullen geven. Bij het gebruik van stationaire satellieten moeten veel grotere afstanden overbrugd worden; deze satellieten bevinden zich 36.000 km boven het aardoppervlak, de totale afstand kan dan ongeveer 80.000 km worden. Met een propagatiesnelheid van 300 km/msec komt men tot een looptijd van 260 msec per richting of 520 msec vertraging. Wordt hierbij de verbinding tussen grondstation en abonnee meegerekend, dan komt men aan een vertraging van ± 600 msec. Een van de gevolgen van de vertragingstijd kan als volgt toegelicht worden: Als A tijdens het spreken even pauseert om B de gelegenheid te geven te reageren, zal deze reactie pas na enige tijd bij A aankomen. Als dit te lang duurt gaat A inmiddels weer verder, de vertraagde reactie van B zal A onderbreken waarna gedurende enige tijd verwarring kan ontstaan. Uit een aantal onderzoekingen (Klemmer '67, Brady '71, Höhne '72) kan afgeleid worden dat de looptijden van 600 - 1200 msec inderdaad verwarring veroorzaken, maar dat dit in het algemeen niet als bijzonder hinderlijk wordt ervaren. De vertraging die optreedt doordat men even op antwoord moet wachten wordt pas bij looptijden groter dan 1200 msec hinderlijk.

III. ECHO-ONDERDRUKKERS

De echo-onderdrukkers die momenteel toegepast worden bestaan in eerste benadering uit een spraakbestuurde schakelaar. Een verbinding met echo-onderdrukkers kan dan weergegeven worden als in fig. 3.

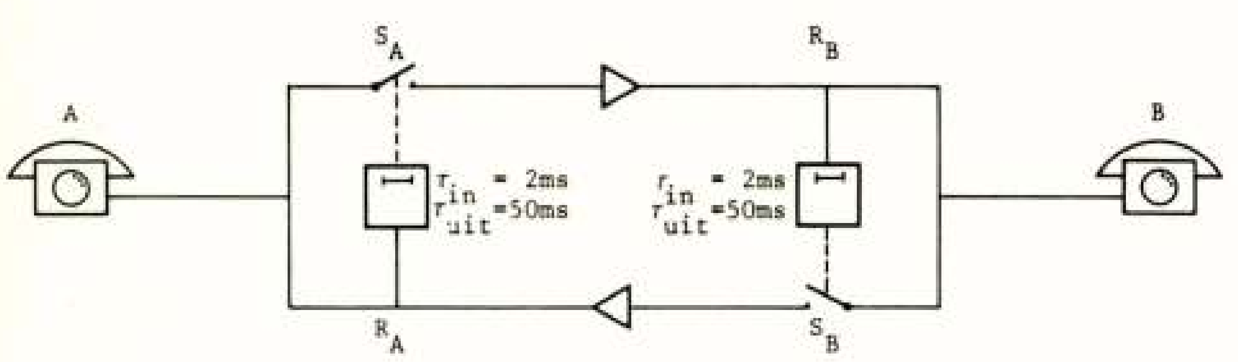


Fig. 3: Vereenvoudigd schema van verbinding met echo-onderdrukker.

De werking is als volgt: wanneer bij R_A spraak gedetecteerd wordt, gaat de schakelaar S_A open zodat de echoweg onderbroken is. Omdat er enige vertraging kan zijn tussen R_A en S_A, blijft S_A nog 50 msec open nadat de spraak bij R_A verdwenen is. Bij de schakeling van fig. 3 is geen interruptie mogelijk, om hierin te voorzien wordt de schakeling uitgebreid volgens fig. 4.

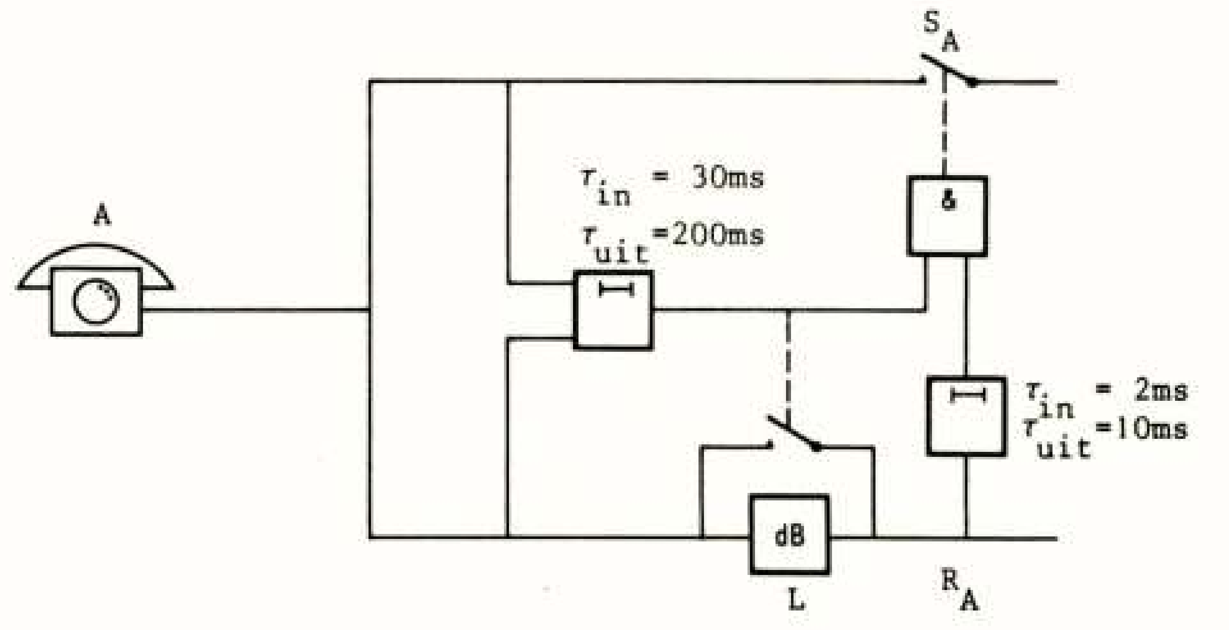


Fig. 4: Principe-schema van echo-onderdrukker.

In deze schakeling is een systeem V opgenomen dat ervoor dient z.g. tweezijdig spreken te detecteren; als dit optreedt wordt de blokkering S_A opgeheven en wordt een demping L aan de R_A zijde ingeschakeld. De bedoeling hiervan is de interruptie van de A-zijde door te laten, maar de echo voor de B-zijde te verzwakken met de demping L.
Het zal duidelijk zijn, dat in deze situatie naar een compromis gezocht is, waarbij de spraak van B verzwakt doorkomt en de echo slechts gedeeltelijk wordt onderdrukt. De gevolgen van dit compromis worden hinderlijker naarmate de looptijd langer is (Karlin '67), zie fig. 5.

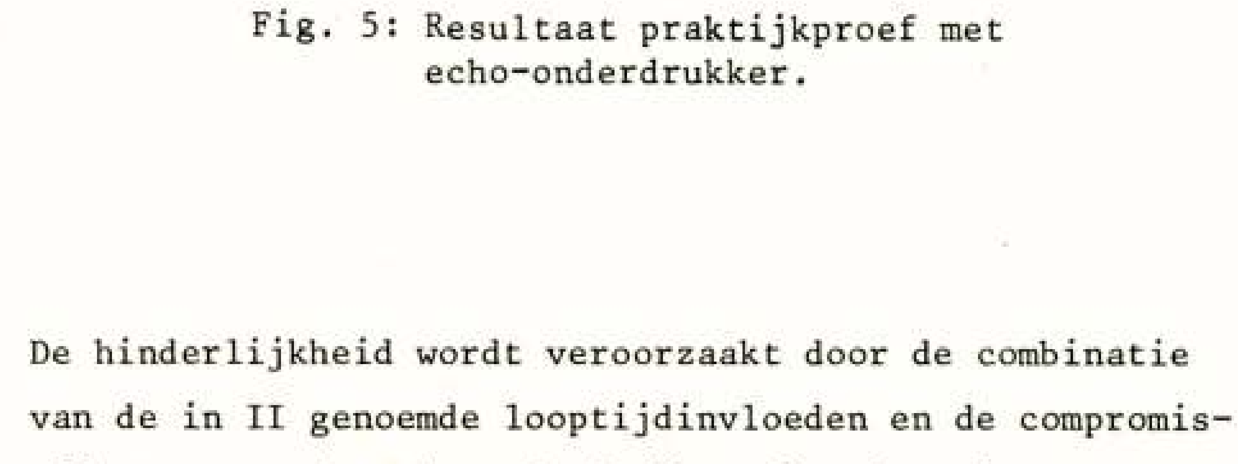


Fig. 5: Resultaat praktijkproef met echo-onderdrukker.

De hinderlijkheid wordt veroorzaakt door de combinatie van de in II genoemde looptijdinvloeden en de compromis-effecten van de echo-onderdrukker. Het functioneren van de echo-onderdrukker wordt bovendien nog ongunstig beinvloed door de grote niveauverschillen die in de praktijk optreden (harde en zachte sprekers, grote of kleine demping).

IV. PRINCIPES VAN DE ADAPTIEVE ECHO-CANCELLER

Bij deze wijze van echobestrijding wordt een model gevormd van de echoweg, zie fig. 6. Het uitgangssignaal van dit model wordt afgetrokken van het werkelijke signaal en met behulp van dit verschilsignaal wordt het model bijgeregeld om de echo te minimaliseren. Op deze manier is het mogelijk de echo te elimineren zonder de problemen, die inherent zijn aan spraakbestuurde schakelaars.

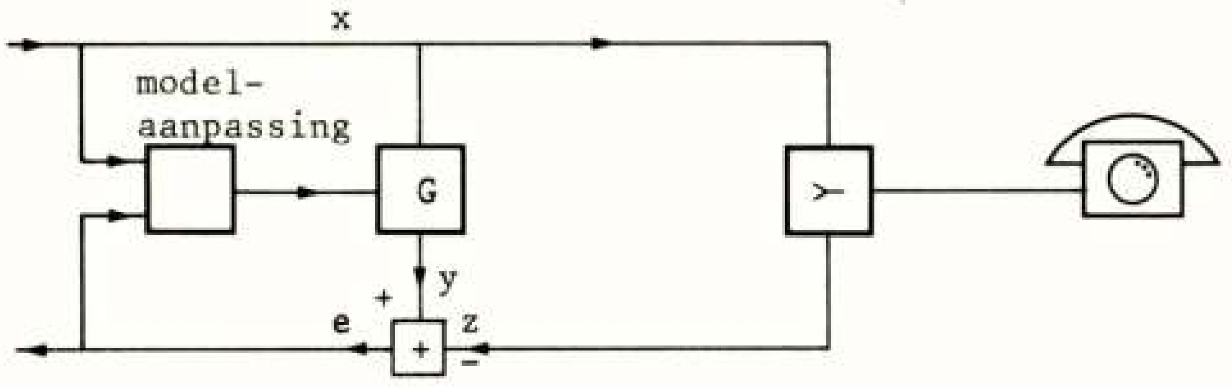


Fig. 6: Principe van adaptieve echo-canceller.

De aanpassing van het model moet gedurende het gesprek plaats blijven vinden, omdat de karakteristieken van de echoweg kunnen veranderen b.v. bij het doorverbinden in een huisautomaat of door een verandering in de akoestische overloop bij het toestel, die afhankelijk is van de stand van de handmicrotelefoon.
Hoewel in theorie verschillende modelconfiguraties mogelijk zijn, is het meest gebruikte model, een transversaal filter, waarvan de weegfactoren gevarieerd worden. Omdat dit op enige schaal en met voldoende nauwkeurigheid alleen digitaal is uit te voeren, kan het model als een discreet transversaal filter worden weergegeven (fig. 7).

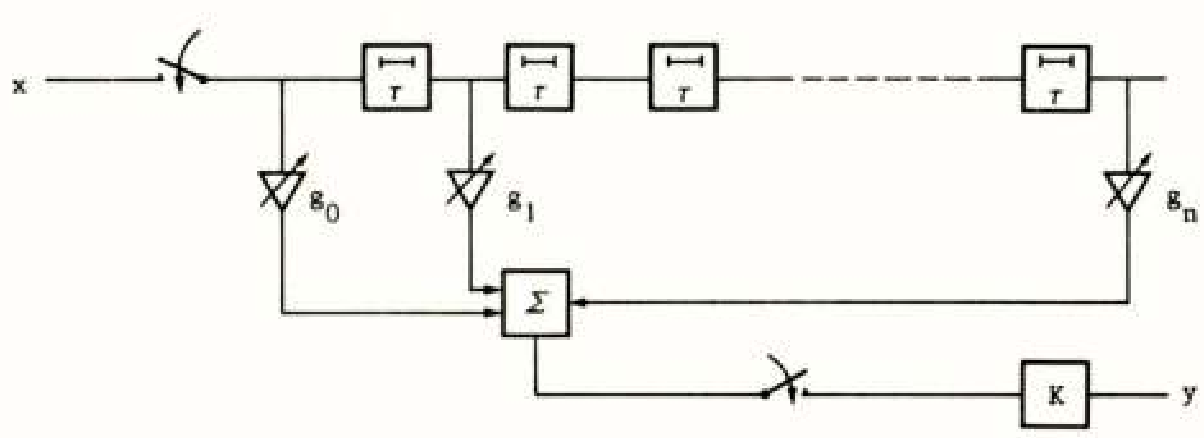


Fig. 7: Digitaal transversaal filter.

Bij dit filter wordt het signaal op discrete tijdstippen aan de ingang bemonsterd en het uitgangssignaal op dezelfde tijdstippen afgegeven. Een filter K verwijdert de ongewenste produkten. De bemonsteringsfrequentie wordt volgens het bemonsteringstheorema gevonden uit de bandbreedte. De gewenste bandbreedte is 3400 Hz, met enige marge wordt dit 4000 Hz, de bemonsteringsfrequentie moet dan 2 x 4000 = 8000 Hz bedragen en $\tau = \frac{1}{8000} = 125$ µsec. De weegfactoren $g_o \dots g_n$ vormen de gediscretiseerde impulsresponsie; het aantal elementen n wordt bepaald door de grootste lengte van de impulsresponsie, die in de echoweg kan optreden. Het gaat hier om de echo in het nationale telefoonnet en de lengte van de impulsresponsie kan voor verschillende landen een verschillende waarde hebben. Voor Nederland zal naar schatting 25 msec voldoende zijn, n is dan $\frac{25}{0,125} = 200$. Om een eerste indruk te krijgen van de werking van de echo-canceller kan de vereenvoudigde voorstelling van fig. 8 dienen. De echo-canceller bestaat uit een transversaal filter met coëfficiënten g_k, deze coëfficiënten worden berekend volgens de formule $g_k = \gamma \int x_k\, edt$. Als rekenmodel voor de echoweg kan een transversaal filter met coëfficiënten h_k dienen. Het uitgangssignaal van de echo-canceller is $y = \sum_{k=1}^{n} x_k\, g_k$, het echosignaal is $z = \sum_{k=1}^{n} x_k\, h_k$. De echo e is volledig gecompenseerd als $g_k = h_k$ voor $k = 1 \dots n$.

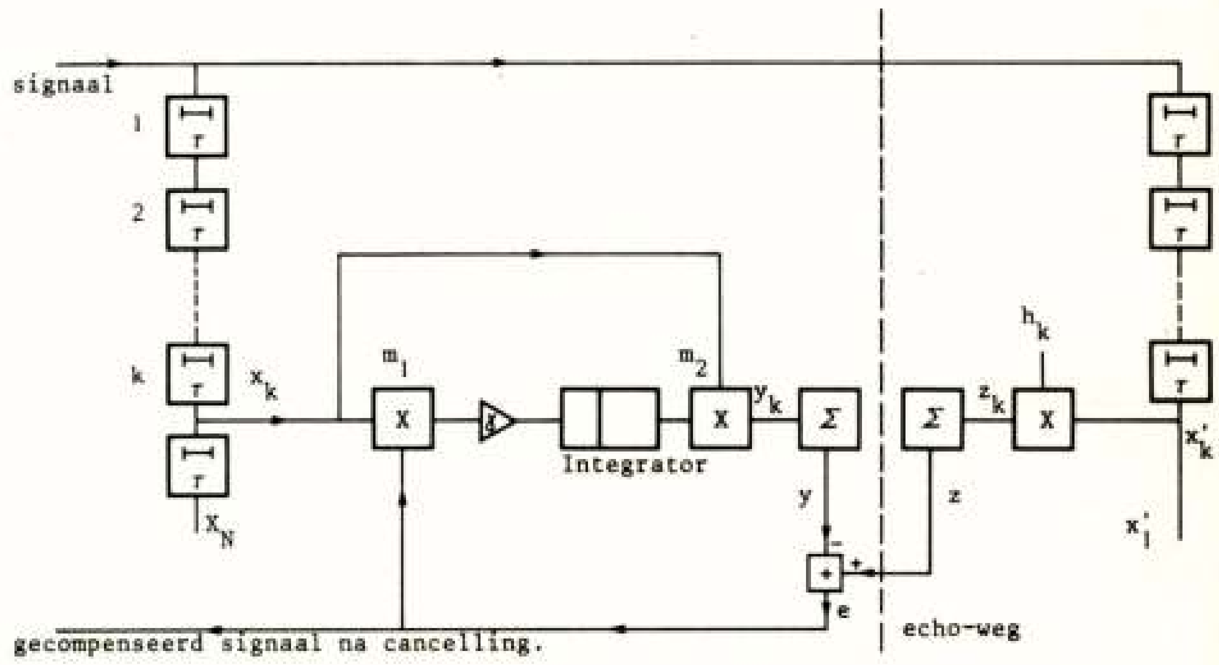


Fig. 8: Vereenvoudigd schema van de echo-canceller met een rekenmodel van de echo-weg.

Wanneer in eerste instantie alleen de componenten y_k en z_k in beschouwing worden genomen, met verwaarlozing van alle andere componenten y_j en z_j ($j = 1 \dots n$) dan is uit fig. 8 gemakkelijk af te leiden dat voor dit deelsysteem g_k convergeert naar h_k. Immers als $g_k > h_k$ en $x_k > 0$ dan is $y_k > z_k$ en $e < 0$ dus $x_k.e < 0$, voor $x_k < 0$ is dan $y_k < z_k$ en $e > 0$ dus $x_k.e < 0$. Dus als $g_k > h_k$ dan is $x_k.e$ altijd negatief, op dezelfde manier is in te zien dat voor $g_k < h_k$ $x_k.e$ positief is. Omdat $g_k = \gamma \int x_k\, edt$ betekent dit dat g_k toeneemt als $g_k < h_k$ en dat g_k afneemt als $g_k > h_k$. Het stabiele punt is $g_k = h_k$ waarbij de echo volledig gecompenseerd is.
In werkelijkheid bevat het e-signaal behalve het verschil van y_k en z_k ook de verschilsignalen van de andere y_j en z_j ($j = 1 \dots n$). Hoewel hierdoor de convergentie wordt vertraagd, kan bewezen worden dat het systeem onder normale omstandigheden convergeert (Sondhi '67).

Bij het echosignaal e moet nog een stoorsignaal n opgeteld worden, dit signaal bestaat uit ruis en spraak van de abonnee, die zich bij de echoweg bevinden. Omdat $g_k = \gamma \int x_k$ (e+n) dt en omdat x ongecorreleerd is met n kan de invloed van n gemiddeld 0 zijn; de momentele invloed van n wordt bepaald door γ. Naarmate γ kleiner is wordt de invloed van n kleiner, de convergentiesnelheid wordt echter ook kleiner zodat een compromis moet worden gezocht.
Een nadeel van het beschreven systeem is, dat de snelheid waarmee g_n varieert evenredig is met het kwadraat van de grootte van het ingangssignaal ($x_k e$). Wanneer x_k en e omgezet worden in een driewaardig signaal volgens fig. 9 treedt dit nadeel niet op, bovendien wordt hierdoor de vermenigvuldiger voor $x_k e$ sterk vereenvoudigd.

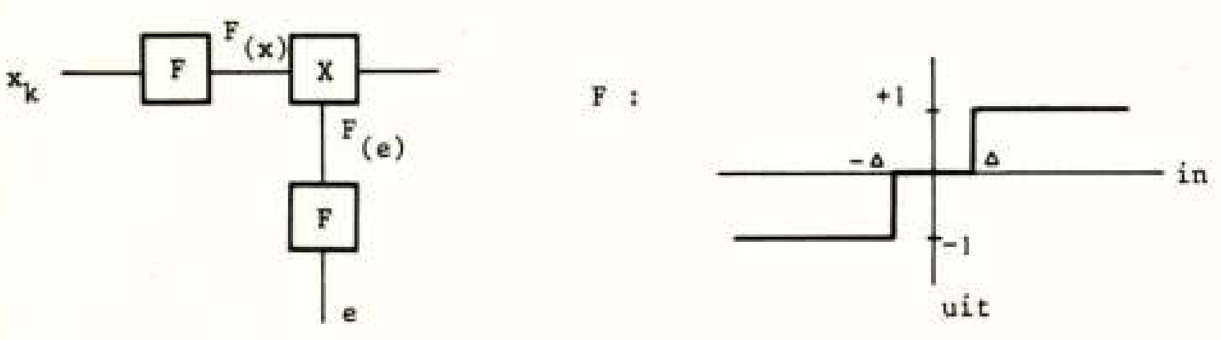


Fig. 9: Vereenvoudiging van de vermenigvuldiger m_1

Voor kleine waarden van x en e is F(x) en F(e) = 0 waardoor de invloed van ruis beperkt wordt; voor grotere waarden is F(x) en F(e) + of -1, de grootte van het signaal is dan niet meer van belang.
Naast bovenstaande gevoelsmatige benadering is ook een meer mathematische beschouwing van de adaptieve echo-canceller mogelijk. Uitgaande van fig. 8 kan geschreven worden:

$$(\tau = 1) \quad y(t) = \sum_{j=1}^{n} x(t-j)\, g_j \; ;$$

$$e(t) = z(t) - \sum_{j=1}^{n} x(t-j)\, g_j$$

Ditzelfde kan in vectornotatie geschreven worden met de vectoren $\phi[t+1] = [x(t), \text{----}\; x(t-n+1)]$

$$\theta(t) = [g_1 \text{ ----- } g_n]^{\tau}$$

$$\rightarrow e(t+1) = z(t+1) - (\phi(t+1).\theta(t))$$

De aanpassing kan nu geschieden door een verliesfunctie $V(\theta(t)) = \int_0^{\tau} (e(t))^2$ te minimaliseren. Dit is met een iteratief proces mogelijk volgens de gradiëntmethode: $\theta(t+1) = \theta(t) - \gamma(t)$ grad $V(\theta(t)$

$$V(\theta t) = [z(t+1) - \phi(t+1)\, \theta(t)]^2$$

$$\text{grad } [V(\theta(t))] = - [\phi(t+1)]^{\tau} [z(t+1) - \phi(t+1)\, \theta(t)]$$

en $\theta(t+1) = \theta(t) + \gamma(t) [\phi(t+1)]^{\tau} [z(t+1) - \phi(t+1)\, \theta(t)]$

Voor een constante waarde van γ komt deze methode overeen met fig. 8.

In de literatuur is een groot aantal methoden bekend om gunstige waarden voor γ te bepalen. In de meest algemene benadering is γ een tijdafhankelijke matrix. Bekend is b.v. de "real time" kleinste kwadratenmethode, die hier ter illustratie wordt gegeven (Åström '71)

$$\theta(t+1) = \theta(t) + \Gamma(t) [\phi(t+1)]^{\tau} [\alpha + \phi(t+1)\, p(t)\, \phi(t+1)^{\tau}]^{-1}$$

$$\Gamma(t) = p(t) [\alpha + \phi(t+1)\, p(t) [\phi(t+1)]^{\tau}]^{-1}$$

$$p(t+1) = [I - \Gamma(t)\, \phi(t+1)]\, p(t)$$

Hierin is α een factor, die bepaalt hoe snel veranderingen in θ kunnen worden gevolgd, waarbij een grotere snelheid gepaard gaat met een grotere gevoeligheid voor ruis. Overigens is het werken met de nxn matrix p in apparatuur moeilijk te realiseren bij waarden van n van $\pm$ 200.

V. PRAKTISCHE REALISATIE

Ingewikkelde algoritmen zoals de in IV vermelde kleinste kwadratenmethode kunnen goed uitgevoerd worden met computersimulaties. Uitvoering in apparatuur is echter moeilijk. Voor de verkenning die het Dr. Neher Laboratorium op dit gebied heeft uitgevoerd is dan ook gekozen voor een eenvoudig proefmodel met een vertragingslijn van 40 stappen. Dit is onvoldoende om echo's in het Nederlands telefoonnet te bestrijden, maar voldoende voor een kennismaking met de techniek van de echo-canceller.

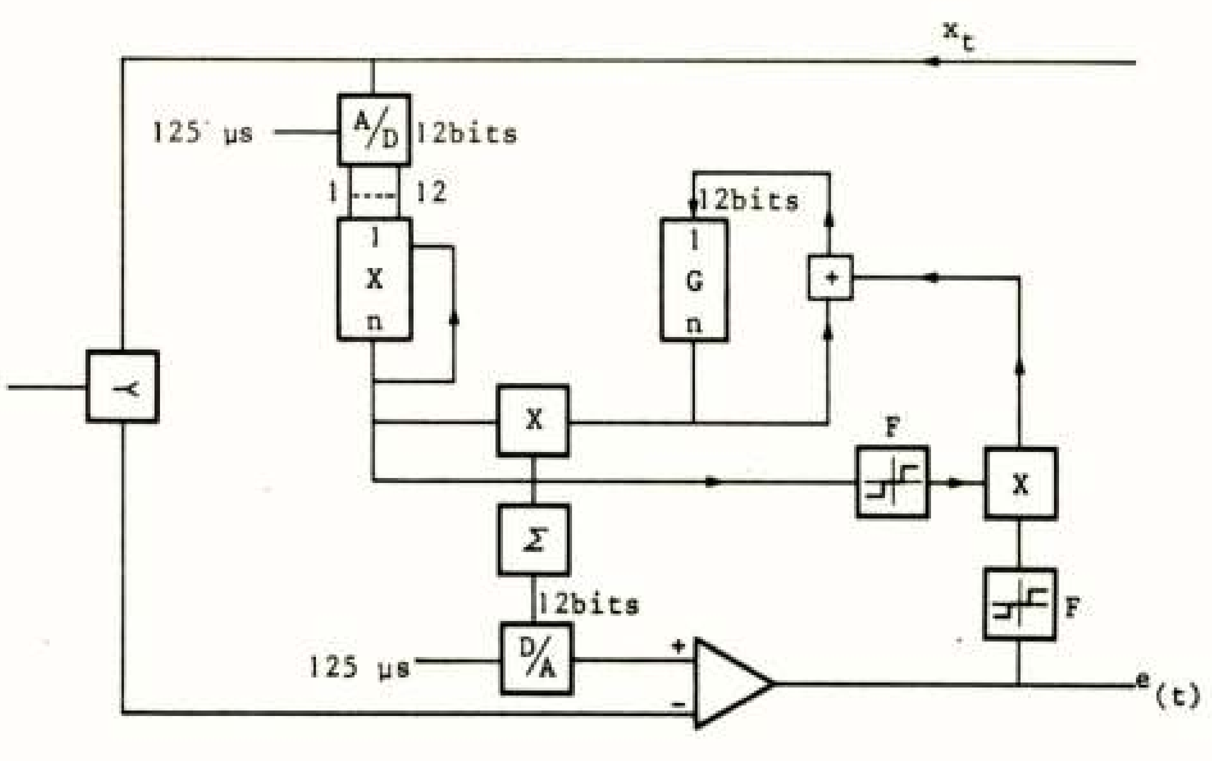


Fig. 10: Blokschema van het DNL - proefmodel.

Het blokschema is gegeven in fig. 10; ten opzichte van fig. 8 zijn de vertragingslijnen vervangen door circulerende schuifregisters X en G. Tussen twee samples in worden beide registers volledig gecirculeerd en worden alle N produkten $G_j\, X_j$ achtereenvolgens berekend en opgeteld. Voor spraak kan met dit model een overloopdemping van $\pm$ 30 dB worden bereikt in 5-10 woorden.

Volgens de literatuur (Campanella '72) kan door een geschikte signaalbewerking een vorkoverloopdeming van meer dan 35 dB worden bereikt. Het restant van de echo kan dan met een z.g. center-clipper voor een groot deel worden verwijderd (fig. 11).

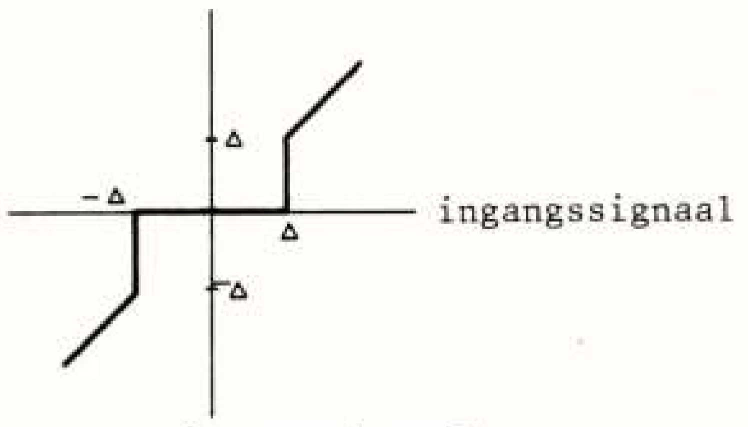


Fig. 11: Center - clipper.

VI. ECHO-CANCELLERS IN DE PRAKTIJK

Praktijkproeven op grote schaal met echo-cancellers zijn uitgevoerd door Intelsat (Suyderhoud '75). Uit de resultaten valt af te leiden dat de waardering van abonnees voor verbindingen met echo-cancellers slechts weinig beter is dan voor verbindingen met echo-onderdrukkers. Voor een deel is dit te wijten aan versluierende effecten zoals ruis en een te grote demping. In de toekomst is te verwachten dat de kwaliteit van telefoonverbindingen verbeterd zal worden. Het gevolg zal zijn dat echoverschijnselen beter merkbaar zijn en steeds meer een beperkende factor voor de kwaliteit zullen vormen. Hierdoor kan de behoefte aan een echo-canceller toenemen.
Over het gebruik van 2 hops-satellietverbindingen met echo-cancellers is nog weinig uit de praktijk bekend. Het is denkbaar dat deze mogelijkheid, het gebruik van echo-cancellers zal stimuleren. Een andere belangrijke factor is de prijsontwikkeling; het gebruik van de momenteel beschikbare IC's leidt tot een groot aantal componenten en een onaantrekkelijke prijs. Voor serieproduktie zou een grotere mate van integratie noodzakelijk zijn.
Een interessante gebruiksmogelijkheid van de echo-canceller is deze te bestemmen als meetinstrument, waarbij men de impulsresponsie van een systeem kan meten zonder gebruik te maken van testsignalen. Op deze wijze kunnen b.v. de karakteristieken van de vorkoverloopdemping worden bepaald zonder de abonnee lastig te vallen.

Literatuur

[1] D.L. Richards - Theoretical study of the functioning of echo-suppressors, Teleteknik '63.
[2] E.T. Klemmer - Subjective evaluation of transmission delay in telephone conversations B.S.T.J., July - August '67.
[3] P.T. Brady - Effects of transmission delay on conversational behaviour on echo free telephone circuits B.S.T.J., January '71.
[4] H.P. Höhne - Influence of transmission delay on testconversations in simulated echo free circuits NTZ-1972 - H11.
[5] I.E. Karlin - Measuring the acceptability of long delay transmission circuits during the Early Bird transatlantic tests in 1965 - Het PTT-bedrijf-XV, mei '67.
[6] M.M. Sondhi - An adaptive echo-canceller B.S.T.J., March '67.
[7] K.J. Åström, P. Eykhoff - System Identification, a survey. Automatika 7 (1971).
[8] S.J. Campanella, H.G. Suyderhoud, M. Onufry - Analysis of an adaptive impulse response echo-canceller. Comsat technical review vol. 2, Spring '72.
[9] M.G. Suyderhoud, S.J. Campanella, M. Onufry - Results and analysis of world wide echo-canceller field trial. Comsat technical review vol. 5, Fall '75.

Voordracht gehouden op 12 mei 1976 tijdens een bijeenkomst van het Nederlands Akoestisch Genootschap, het Nederlands Electronica- en Radiogenootschap en de Benelux sectie IEEE in het Instituut voor zintuigfysiologie TNO te Soesterberg.

L.P.C.-ANALYSE EN FORMANTSYNTHESE VAN SPRAAK

Ir. L. F. Willems
Instituut voor Perceptie Onderzoek, Eindhoven

In het kort worden hier de principes van een betrekkelijk nieuwe analyse-synthese-methode van spraak beschreven, met welke methode een goede kwaliteit van de gereproduceerde spraak is te bereiken (Linear Predictive Coding). Ten behoeve van het fonetisch onderzoek geschiedt het synthetiseren met behulp van formanten.

1) INLEIDING

Het vinden van een representatie van spraaksignalen in een aantal slechts langzaam variërende parameters is van groot belang voor een aantal toepassingen in het spraakonderzoek. De mens produceert een gering aantal (5 à 10) spraakklanken per seconde, want de articulatoren en dus ook het spraakkanaal kunnen slechts met een beperkte snelheid bewegen. Ook het menselijk gehoororgaan, langs welke weg de spraak bij de mens binnenkomt , is beperkt wat betreft het verwerken van het aantal verschillende klanken per seconde. Uit dit soort van overwegingen is het plausibel te maken dat de informatie-inhoud van bijvoorbeeld PCM-gecodeerde spraak van 48000 bits/sec (6000 Hz aftastfrequentie en 8 bits per sample) aanzienlijk te reduceren is. Als ondergrond komt men tot ongeveer 60 bits/sec. (stelt men dat er 64 verschillende spraakklanken zijn, die in 6 bits zijn te coderen, en stelt men dat 10 spraakklanken per seconde worden geproduceerd, dan geeft dit 60 bits/sec.).

Toepassingen van deze in de praktijk echter niet zo dramatische informatie-reductie bij spraaksignalen, liggen op het terrein van de spraakherkenning, van spraakopslag t.b.v. zogenaamde voice response en van de overdracht van spraak. Op het gebied van de spraaktransmissie noemt men deze toepassingen: vocoders. Hoewel het idee van de vocoder uit de dertiger jaren stamt en in de tijd daarna er veel uitvoeringsvormen van vocoders zijn voorgesteld (SCHRÖDER 1966), zijn ze nooit op grote schaal toegepast, ofwel de bereikte bandbreedtereductie was te gering ofwel de kwaliteit van het uiteindelijke spraaksignaal was te slecht. De laatste jaren is de belangstelling voor vocoders weer toegenomen, voornamelijk door de grote vlucht van de digitale technieken. Ook biedt de digitale vorm goede perspectieven voor geheime coderingen tijdens de spraakoverdracht.

Een van de meer complexe, maar toch in de praktijk uitvoerbare analyse-synthese-technieken is de zog. Linear Predictive Coding: LPC afgekort. (ITAKURA 1969, ATAL and HANAUER 1971). In het nabije verleden is aangetoond dat enerzijds zeer goede kwaliteit van de spraak is te bereiken (ATAL) en anderzijds een grote bandbreedtereductie tot beneden 1000 bits/sec is te behalen (SAMBUR 1975, KANG and COULTER 1970).

De motivatie voor ons om aan dergelijke analyse-synthese-systemen te werken zijn de toepassingen ervan bij het fonetisch onderzoek. Bij studies omtrent de waarneming van spraakklanken heeft de experimentator vaak behoefte aan stimuli, die volgens bepaalde voorschriften zijn gemaakt of gevarieerd. Het maken van dergelijke stimuli kan geschieden m.b.v. analyse-synthese-systemen van spraak. Wil men de waarneming van de intonatie van zinnen bestuderen, dan moet men de luisteraar (proefpersoon) zinnen kunnen voorspelen, waarin de intonatie systematisch wordt gevarieerd. Daartoe wordt de toonhoogte (dat is de grondfrequentie van het brongeluid) bij het syntheseproces vervangen door een kunstmatig opgewekt toonhoogteverloop, zoals dat door de experimentator wordt gevraagd.

In deze bijdrage wordt een korte beschrijving gegevens van deze LPC analyse-synthese-techniek. In de fonetiek is echter een beschrijving van spraakklanken in termen van de zog. formanten gebruikelijk. Formanten zijn de resonantiefrequenties van het mondkanaal. Bij een neutrale klinker van een mannenstem liggen de formanten resp. bij: 500 Hz eerste formant, 1500 Hz tweede formant, 2500 Hz derde formant. De LPC techniek levert een goede mogelijkheid de analysegegevens om te rekenen naar een formantenbeschrijving. Zodoende kan het syntheseproces gebruik maken van deze in de fonetiek gekende formanten.

2) LINEAR PREDICTIVE CODING

Hierbij gaat men uit van een productiemodel van spraak dat bestaat uit een bron, die ofwel een periodieke puls ofwel ruis produceert en een lineair filter dat alleen polen bevat (fig.1a).

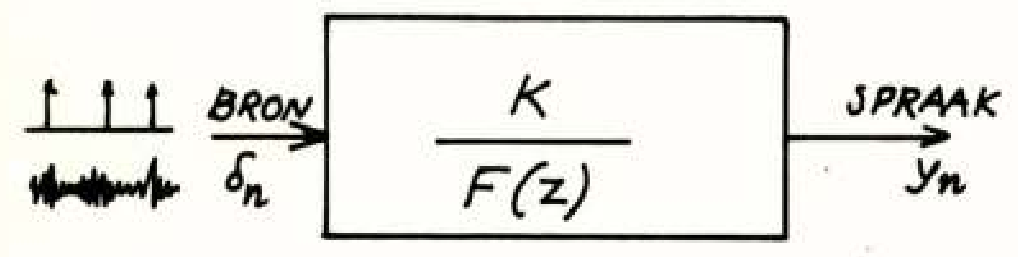


Fig. 1a. Spraakproductiemodel.

Hoewel de overdrachtsfunctie van het mondkanaal in sommige gevallen ook nulpunten bevat en hoewel het brongeluid, dat door de stembanden wordt geproduceerd, niet pulsvormig van aard is, is het de veronderstelling dat met het genoemde model het spraakproductieproces voldoende nauwkeurig is te benaderen. Het filter dat alleen polen bevat is in fig. 1b voorgesteld als een recursief filter:

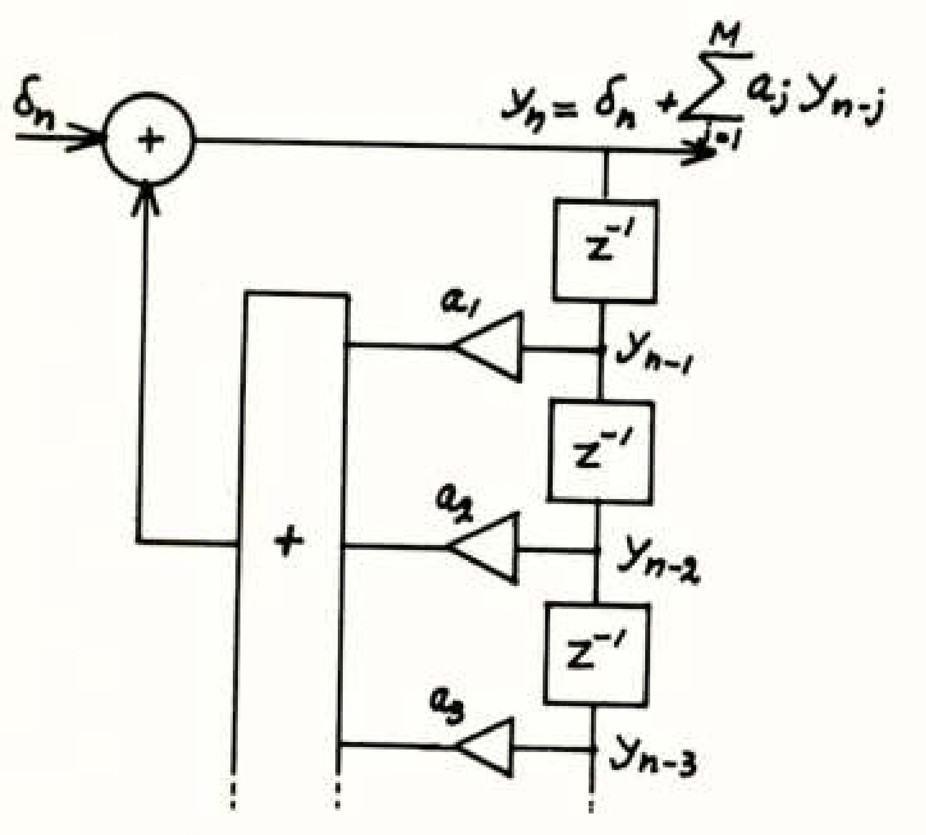


Fig. 1b. Spraakproductiemodel als recursief filter.

ook output spraaksample S_n kan worden geschreven als de som van de input δ_n en een lineaire combinatie van een aantal spraaksamples S_{n-j} uit het verleden:

$$S_n = \delta_n + \sum_{j=1}^{m} a_j S_{n-j} \qquad (1)$$

De lineaire combinatie van een aantal spraak samples uit het verleden is op te vatten als een voorspelling; vandaar de naam Linear Predictive Coding. De parameter M bepaalt de orde van het filter en is het aantal polen dat bij de benadering wordt gebruikt. M ligt in de praktijk tussen 10 en 15. De coëfficienten a_j voor $5 = 1, 2, \ldots, M$ bepalen de overdrachtsfuncties en ze vormen samen met nog enkele andere parameters (Fo en stemhebbend- stemloos parameter) een beschrijving van het spraaksignaal op een bepaald moment. De waarden van deze coëfficienten a_j veranderen betrekkelijk langzaam en worden in de praktijk 100 of 50 per seconde bepaald.

Bij de bepaling van de coëfficienten a_j gaat men uit van een stuk spraak van bijvoorbeeld 25 ms, overeenkomend met 250 samples als met 10 kHz wordt afgetast. Dit aantal noemen we N. De veronderstelling is dat het spraakkanaal gedurende dat tijdsinterval als stationair is te beschouwen. Het stelsel coëfficienten $\{a_j\}$ wordt uit dit stuk spraak door een minimaliseringsprocedure bepaald. Dit geschiedt als volgt: Door de predictor (het filter) kan een spraaksample worden voorspeld. Daarbij treedt een fout op t.o.v. het spraaksample S_n in het analyse-interval:

$$E_n = S_n - \sum_{j=1}^{M} a_j S_{n-j} \qquad (2)$$

Door het minimaliseren van de gemiddelde kwadratische fout: $\{E_n^2\}$ gem worden de coëfficienten $\{a_j\}$ gevonden.

$$\frac{\partial}{\partial a_k}\{(S_n - \sum_{j=1}^{M} a_j S_{n-j})^2\}_{gem} = 0 \qquad (3)$$

voor K = 1,2,.....,M.

Hieruit volgt een stelsel vergelijkingen:

$$\sum_{j=1} a_j R_{|K-j|} = R_K \text{ voor } K=1,2,\ldots,M \qquad (4)$$

waarin

$$R_K = \sum_{n=0}^{N-1-k} S_n S_{n+k} \qquad (5)$$

Deze laatste grootheden R_k zijn de auto-correlatie-coëfficienten van het stuk spraak in het analyse interval. De matrix die in dit stelsel vergelijkingen voorkomt is van een speciale vorm, waardoor het stelsel vergelijkingen recursief en snel is op te lossen.

Aan de parameters nodig om het spraaksignaal compleet te beschrijven ontbreken nog enkele (fig. 1a), nl. de amplitude van het signaal, het gegeven of ruis of periodiek signaal als bron moet dienst doen en in het geval het bronsignaal periodiek is, is het nodig de herhalingsfrequentie ervan te kennen. Het meten van deze herhalingsfrequentie is een probleem waarop we later nog terugkomen.

De voorspellen (het filter) die bij de bepaling van de coëfficienten aj wordt gehanteerd wordt ook wel invers filter genoemd (fig. 2). De output van het filter als men de spraak op de input zet is immers (in de zin der kleinste kwadraten) geminimaliseerd. Het inverse filter is

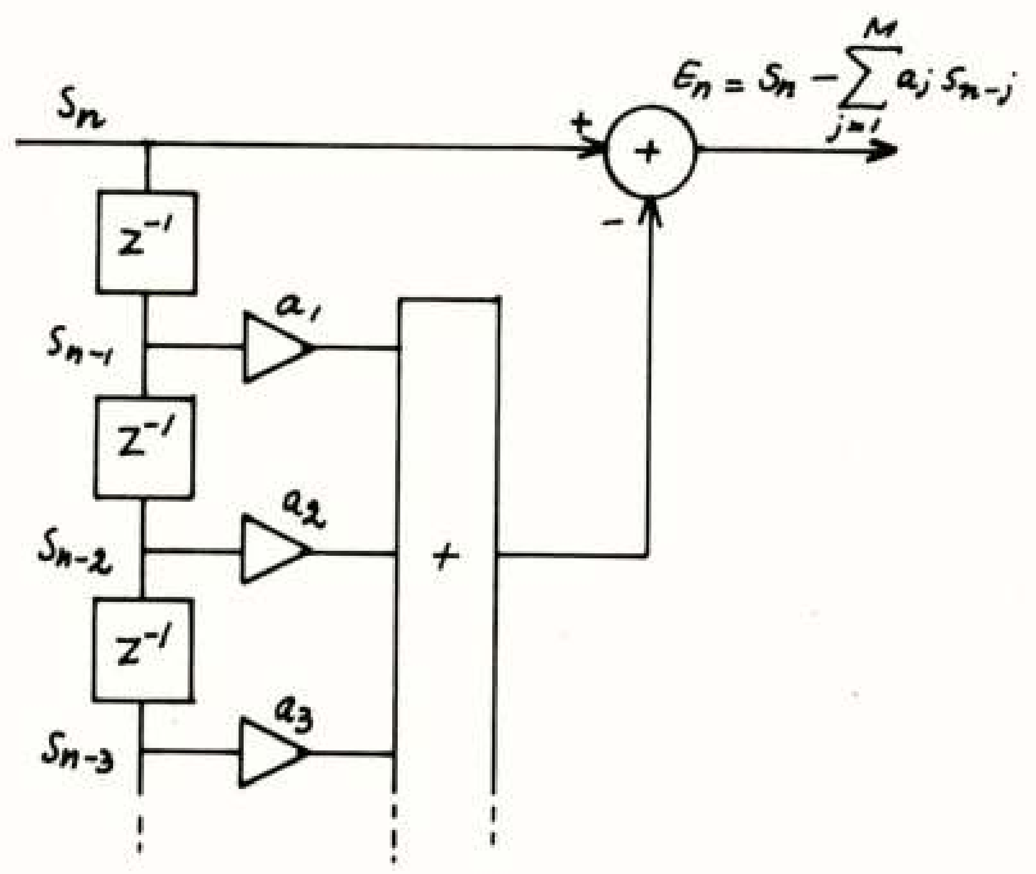


Fig. 2. Voorspel-fout bij de analyse.

dan ook op te vallen als een bewerking op het ingangssignaal welke naar vermogen (aantal polen M) het spectrum probeert glad te strijken.

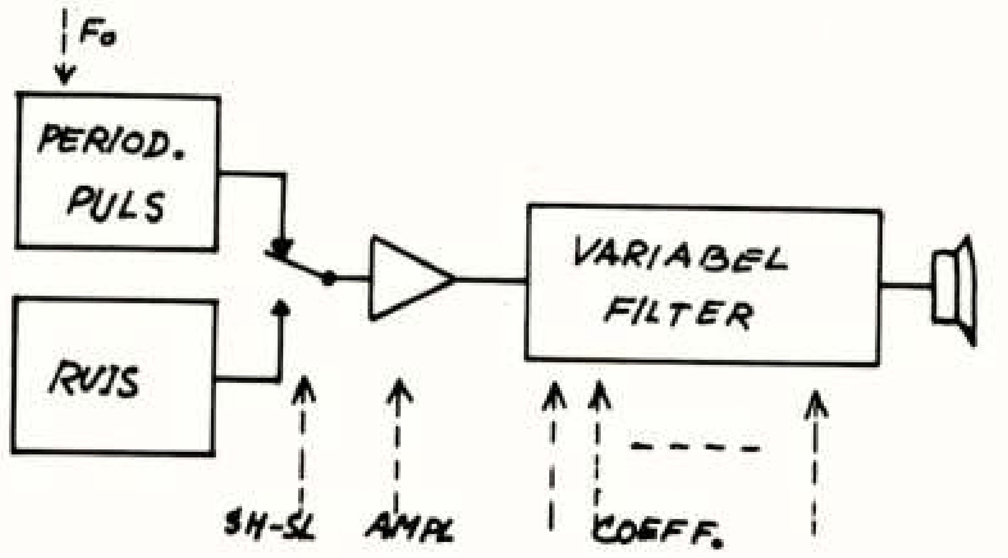


Fig. 3. Synthese-schema.

Het schema dat wordt gebruikt om de spraak m.b.v. deze beschrijving weer te resynthetiseren (fig. 3) correspondeert vanzelfsprekend met het spraakproduktiemodel, dat ten grondslag ligt aan deze methode: De stemhebbend- stemloosschakelaar laat ofwel ruis ofwel een periodieke puls met herhalingsfrequentie Fo toe tot de amplitude modulator en daarna aan het variable filter.

De structuur van het varíabelefilter kan allerlei vormen aannemen. De eenvoudigste structuur is een recursief filter met de coëfficienten a_j zoals in figuur 1b. Dat is een goede manier voor de synthese van spraak en levert goede kwaliteit van de spraak (Atal). Men kan het recursieve filter schrijven als:

$$G(z) = \frac{1}{1+a_1z^{-1}+a_2z^{-2}+a_3z^{-3}+\ldots+a_Mz^{-M}} \quad (6)$$

Het nadeel van deze methode voor ons is, zoals in de inleiding werd aangeduid, dat ze niet aansluit bij de formantstructuur die bij fonetici is ingeburgd. Veel spraakperceptie experimenten zijn gedaan met formanten als variabele grootheden. Bovendien is bekend dat een formantbeschrijving van spraak de zuinigste is wat betreft informatie-inhoud. Een derde reden om naar een formantbeschrijving te zoeken is het feit dat er in ons laboratorium een digitale hardware formantensynthetisator aanwezig is, waarmee synthese in real-time mogelijk is. Daardoor krijgt de experimentator een snelle respons van het systeem, hetgeen enorme voordelen heeft.

De overdrachtfunktie van een in serie geschakelde formanten synthetisator kan men schrijven als: (met 5 formanten)

$$G(z) = \frac{1}{(1+p_1z^{-1}+q_1z^{-2})(1+p_2z^{-1}+q_2z^{-2})\ldots} \quad (7)$$

Willen we de spraak m.b.v. formanten resyntiseren, hetgeen we ons ten doel gesteld hebben, dan moeten we het resultaat uit vergelijkingen (6) nog omrekenen naar p,q - data volgens verg (7). Voor dit doel zijn procedures bekend om kwadratische termen van een polynoom af te splitsen.

3. HET ANALYSEPROGRAMMA

Het computerprogramma, dat voor het berekenen van het inverse filter is ontwikkeld, ziet er in grote trekken als volgt uit.

Het spraaksignaal is op de gebruikelijke wijze gedigitaliseerd. (laagdoorlaat-filter, sample and hold-schakelaar en conversie door een 12 bits ADC), en is opgeslagen in het geheugen van de computer. De sample-frequentie bedraagt 10 kHz. Voor een analyseslag worden 250 samples genomen, overeenkomend met 25 ms. Allereerst wordt het aantal nuldoorgangen in het analyse-interval geteld. Dit gegeven wordt gebruikt voor de beslissing of het signaal stemhebbend (periodiek brongeluid gepaard gaand met een gering aantal nuldoorgangen) ofwel stemloos is (ruis als brongeluid en groot aantal nuldoorgangen). Vervolgens wordt het signaal geschaald en dan wordt preemphasis toegepast door een filter:

$$P(z) = 1 - \mu z^{-1}$$

De constante μ wordt voor een analyse interval en is

$$\mu = \frac{R_1}{R_0}$$

Hierin zijn R_1 en R_0 de eerste en nulde autocorrelatiecoëfficienten. Als de spraak stemhebbend is, dan is μ ongeveer 1 en het preemphasis filter is dan een differentiatie. Bij stemloze klanken kan μ gelijk -1 worden en dan is het preemphasis filter gelijk aan een integrator. Het preemphasis filter moet er voor zorgen dat toppen die in het spectrum voorkomen

ongeveer even hoog komen te liggen. Het preemphasis filter is dan ook op te vatten als een invers filter van de eerste orde. Na deze bewerking wordt het stuk spraak vermeningvuldigd met een hammingwindow

$$X_n = S_n\ (.54 - .46 \cos \frac{2\pi n}{N-1})\ ,\ \text{met } n=1....N.$$

Deze bewerking is gebruikelijk om de ongewenste effecten van een rechthoekig venster te vermijden. Dan volgt het berekenen van het inverse filter met M=10. Daartoe worden 10 autocorrelatie coëfficienten bepaald en wordt het stelsel vergelijkingen (6) opgelost om de filter coëfficienten a_k te vinden. Hiervoor is een snel en recursief algorithme ontwikkeld (MULLER 1973). Zoals in paragraaf 2 is besproken wordt nu dit gevonden filter (van de vormen (6)) in 5 kwadratische termen gesplitst, m.b.v. het Bairstow algorithme. Dit is een iteratieve procedure, die in een enkel geval geen kwadratische term kan vinden. In zo'n geval wordt de term, die bij het vorige analyse interval is gevonden daarvoor in de plaats gezet. Een veel belangrijker probleem is dat de polen niet steeds in een vaste volgorde worden gevonden, terwijl de formanten wel een ordening hebben. Daarom moeten de polen aan de formanten worden toegewezen.

4. HET METEN VAN DE TOONHOOGTE.

Het meten van de toonhoogte is ook een probleem waarvoor sinds decennia naar een oplossing wordt gezocht. Geen van de voorgestelde methoden werkt zonder fouten: de ene methode is zeer gevoelig voor de kwaliteit van het inputsignaal, de andere methode geeft fouten bij kleine signalen en een derde is niet bestand tegen signalen met een sterke tweede harmonische van de grondtoon, enz.

Wij hebben een methode (aangevuld met nog een enkel idee) uit de literatuur (SONDHI 1968) genomen, waarvan bekend is dat hij redelijk betrouwbaar werkt. Voor het meten van de toonhoogte wordt 35 ms (=350 samples) aan het ingangssignaal genomen, om er zeker van te zijn dat minstens twee periodes in het analyse-interval liggen. Dit signaal wordt (niet gewindowed) center-geclipt en vervolgens wordt de autocorrelatiefunctie bepaald. De autocorrelatiefunctie wordt alleen berekend rondom de periode die wordt verwacht. Het maximum in de autocorrelatiefunctie wordt aangewezen als de periode in het ingangssignaal. De grootte van het interval waarin naar het maximum wordt gezocht hangt af van de hoogte van de top in het vorige analyse-interval. Is de top hoog geweest dan is dat een duidelijke periode geweest en dan wordt de breedte van dit venster smal gezet. Deze methode bespaart niet alleen rekentijd maar houdt ook rekening met een zekere continuïteit die in het verloop van de natuurlijke toonhoogte aanwezig is.

5. SLOT OPMERKING

Een korte inleiding werd gegeven van een analyse synthese-systeem gebaseerd op Linear Predictive Coding. Naast de genoemde eigenschap, dat de kwaliteit van de gereproduceerde spraak goed is zijn nog meer eigenschappen interessant die hier onbesproken moesten blijven. We noemen slechts:

- De analyse is bestand tegen storing in het ingangssignaal: het minimaliseringsproces verloopt normaal, zij het met een iets grotere restfout.
- De hier geschetste analysemethode levert een een filterpolynoom die altijd stabiel is.
- Het filter is ook te schrijven als ladderstructuur, die equivalent is aan het mondkanaal opgevat als akoestische buis.

De theorie is in korte tijd (minder dan 10 jaar) gegoeid tot een methode, die op bijna alle gebieden van het spraakonderzoek is doorgedrongen als een gevestigd gereedschap (MARKEL & GRAY 1970).

LITERATUUR

Schroeder, M.R. (1966) Vocoders, Analysis and Synthesis of Speech

Itakura, F. en Saiso, S. (1969) Speech Analysis-Synthesis System based on the partial Autocorrelation Coefficient. Acoust. Soc. of Japan Meeting 1969.

Atal, B.S. en Hanauer, S.L. (1971) Speech Analysis and Synthesis by Linear Predictive of the Speech Wave. J.A.S.A. 50, 1971. 637-655

Sambur, M.R. (1975) An Efficient Linear Prediction Vocoder. BSTJ vol. 54, 1975 , pp. 1693-1723.

Kang, G.S. en Coulter D.C. (1976) 600 BPS Voice Digitizer Int. Conf. ASSP 1976.

Muller, H.F.(1973) Eem methode voor het oplossen van een stelsel vergelijkingen, met een symmetrische coefficienten matrix, IPO Memorandum 122.

Sondhi, M.M. (1968) New Methods of Pitch Extraction IEEE trans. on Audio vol Av-16, 1968, pp. 262-266.

Markel, S.D. en Gray, A.H. jr. (1975) Linear Prediction of Speech.Springer 1976.

Voordracht gehouden op 12 mei 1976 in het Instituut voor Zintuigfysiologie TNO te Soesterberg op een gemeenschappelijke vergadering van het NERG (no. 256), de Benelux-section IEEE en het Nederlands Akoestisch Genootschap.

ANALYSE EN SYNTHESE VAN SPRAAK MIDDELS EEN BANDFILTERSYSTEEM

Ir. Louis C.W. Pols
Instituut voor Zintuigfysiologie TNO, Soesterberg

For the real-time analysis of speech we make use of a parallel set of bandfilters, which outputs are sampled every 10 msec. The raw spectral data are stored on the disk memory of a computer and can be used for subsequent data processing.

However, these data can also be used for a real-time speech synthesis. The synthesis system, also based on a parallel set of bandfilters, gets its data from the computer and all synthesis parameters can be modified under program control.

In this way not just the originally analyzed speech utterances can be regenerated with a good intelligibility, but also any wanted part out of the text can be isolated and/or repeated for careful listening. If wished so, pitch, loudness, or pause duration can be modified.

This system is very appropriate for generating stimuli for speech perception experiments, and is intensively used for studying vowel coarticulation effects.

INLEIDING

In relatie tot het thema van deze serie voordrachten, namelijk de "transmissie en synthese van spraak", wordt bij analyse en synthese van spraak wellicht in eerste instantie gedacht aan de beschrijving van een systeem voor spraaktransmissie, al dan niet met een lage bitrate of bandbreedte. Er zijn echter nog andere redenen waarom men spraak zou willen analyseren en/of synthetiseren.

Men kan spraak *analyseren* om de fundamentele eigenschappen van het spraaksignaal en zijn dynamische variaties te onderzoeken, of om te komen tot automatische spraakherkenning, of spreker herkenning of -verifikatie, of men kan de verkregen informatie in een of andere vorm visueel of taktiel presenteren voor het spraakonderricht aan doven.

Zonder voorafgaande analyse is het toch mogelijk spraak te *synthetiseren* via synthese door regels. Een dergelijk programmeerbaar synthesesysteem kan in het algemeen gebruikt worden voor signaalgeneratie, niet alleen in de vorm van spraak maar eventueel b.v. ook muziek.

Analyse plus synthese van spraak worden gekombineerd in spraaktransmissie systemen, al dan niet inklusief spraakkodering. In het fundamenteel spraakonderzoek kan een dergelijk systeem echter ook met vrucht gebruikt worden voor de (evt. gedeeltelijke) hersynthese van een stuk spraak, waarbij dan al of niet een of meerdere parameters kunnen worden gemodificeerd. Men kan hierbij denken aan intonatie- en koartikulatieonderzoek. Ook kunnen zo stimuli voor perceptieve proeven worden gegenereerd.

Ten behoeve van het experimenteel spraakonderzoek op het I.Z.F. beschikten wij ook graag over een analyse-synthese systeem dat aan de volgende eisen zou moeten voldoen: De synthetische spraak moest van een redelijke kwaliteit zijn, het systeem moest automatisch zijn en onder programmakontrôle kunnen worden gemanipuleerd. Het systeem moest liefst in real time werken en bij voorkeur gebaseerd zijn op een principe representatief voor de spraakperceptie.

Onze keuze is hierbij gevallen op een kanaalvocoder-achtige benadering. Andere mogelijkheden vielen af door moeilijk te objektiveren analysemethodes (formantvocoder), of door de komplexiteit en de te grote computerbehoefte (LPC-vocoders).

In de volgende paragraaf zal het systeem worden beschreven waarna enkele toepassingen van het systeem fragmentarisch zullen worden besproken.

Beschrijving van het analyse-synthese systeem

Zowel aan de analyse- als aan de synthesekant wordt gebruik gemaakt van een parallelle serie bandfilters. In plaats van een transmissieweg fungeert de computer als intermediaire opslag en aktieve datamanipulator. Veel variatie is mogelijk in het type filter, de onderlinge afstand op de frekwentieschaal, de bandbreedte e.d. Wij hebben ons bij onze keuze laten leiden door het beperkte frekwentieoplossend vermogen van het menselijk gehoororgaan, uitgedrukt in de zogenaamde kritische bandbreedte (Plomp, 1976). Tertsfilters hebben een bandbreedte die goed hiermee overeenkomt. Van 400 Hz tot en met 8000 Hz hebben we hier dan ook gebruik van gemaakt. Teneinde de konstante kritische bandbreedte beneden ca. 400 Hz te simuleren zijn nog drie 90-Hz filters met middenfrekwenties van 122, 215 en 307 Hz aan het systeem toegevoegd. Dit

resulteert in 17 filters.

Analyse
Aan de analysekant is aan deze filters nog een breedbandig filter (-3 dB punten bij 32 Hz en 8000 Hz) toegevoegd om het overall nivo van de spraak te kunnen meten. Deze filters worden gevolgd door logarithmische versterkers en omhullende piekdetektoren waarvan de parallelle uitgangen 100 keer per sek. worden bemonsterd met een 20-kanaals multiplexer. Deze informatie gaat via een analoog-digitaal omzetter naar een PDP-15 computer. Voor een blokschema van het analysesysteem zie Fig. 1.

kunnen geven is gebruik gemaakt van een datareduktie techniek. Hiertoe worden de 17 filterwaarden per 10-msek sample beschouwd als de koordinaatwaarden van een punt in een 17-dimensionale ruimte. Ieder bandfilterspektrum is dan een punt in die ruimte, en opeenvolgende spektra, zoals in een woord, vormen een spoor in die ruimte. Met behulp van een principale-komponenten analyse (v.d. Geer, 1967) kan nu een laag-dimensionale subruimte worden gedefinieerd waarin niettemin een zeer groot deel van de variatie in de oorspronkelijke data kan worden beschreven. Voor display doeleinden is een twee-dimensionale weergave uiteraard het meest geeigend.

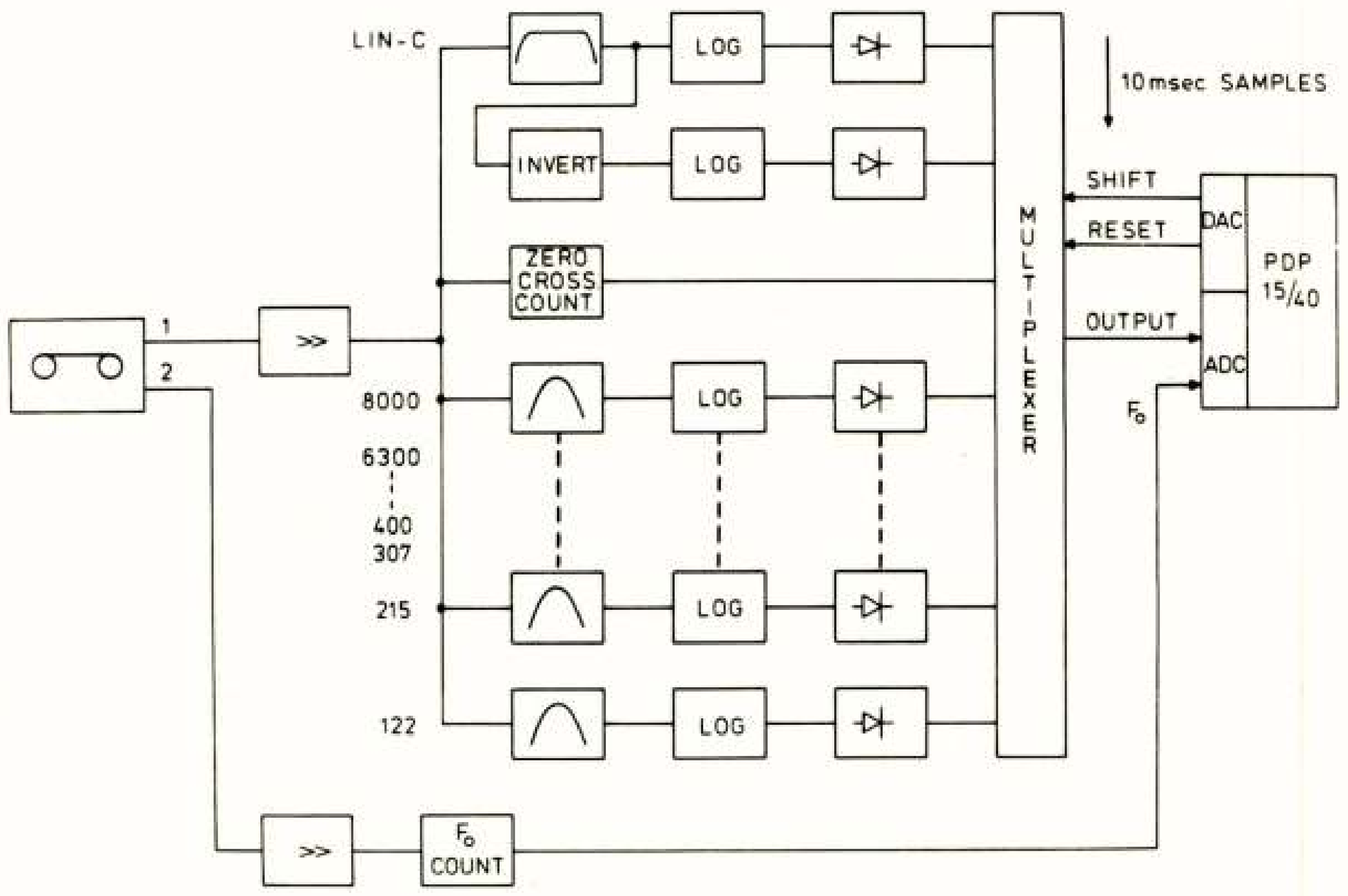


Fig. 1. Blokschema van het analysesysteem.

De ruwe data worden opgeslagen op een schijfgeheugen. Via diverse verwerkings- en displayprogramma's kunnen deze data dan worden bestudeerd. De data kunnen, na eventuele modifikatie, ook gebruikt worden voor een gehele of gedeeltelijke hersynthese. Zowel analyse als synthese gebeuren in real time, de tussenbewerkingen zijn echter off-line. Alvorens het synthesegedeelte te bespreken zullen we enige voorbeelden geven van de verwerkings- en displaymogelijkheden.

Dataverwerking en parameterdisplay
De ruwe gegevens per 10-msek sample kunnen allereerst in de vorm van een getallendisplay worden weergegeven. Het verloop van iedere parameter kan ook als funktie van de tijd zichtbaar gemaakt worden. Fig. 2 is een voorbeeld van een dergelijke display.

Door een pointer over het scherm te bewegen worden de numerieke waarden op ieder moment aangegeven in het rechter gedeelte van de display. Rechtsboven is het bandfilterspektrum op het betreffende tijdmoment zichtbaar. Om het *spektrale* verloop als funktie van de tijd weer te

Fig. 3 geeft ter illustratie hiervan een weergave van de gemiddelde positie in de twee-dimensionale subruimte van een aantal klinkersegmenten geisoleerd uit eenlettergrepige woorden. Als klinkers zijn gekozen de *ie*, *oe*, en *aa*. Een dergelijke weergave vertoont veel verwantschap met die in het zogenaamde formantvlak, is echter ten opzichte van deze eenduidiger en, wat zeer belangrijk is, kan automatisch worden bepaald.

Deze benadering biedt tevens de mogelijkheid tot een visuele representatie van de spektrale informatie ten behoeve van het spraakonderricht aan doven. Een eenvoudiger versie hiervan is in de praktijk met sukses getoetst (Povel, 1974). Vervolgens wil ik in het kort het synthese systeem beschrijven.

Synthese
Uit de literatuur is bekend (Flanagan, 1972) en beide eerste voordrachten hebben dat opnieuw aangetoond (Sluyter; Kuijk en Franssen, dit nummer), dat een voice-excited kanaalvocoder goede kwaliteit spraak kan produceren. Deze benadering was echter voor ons type on-

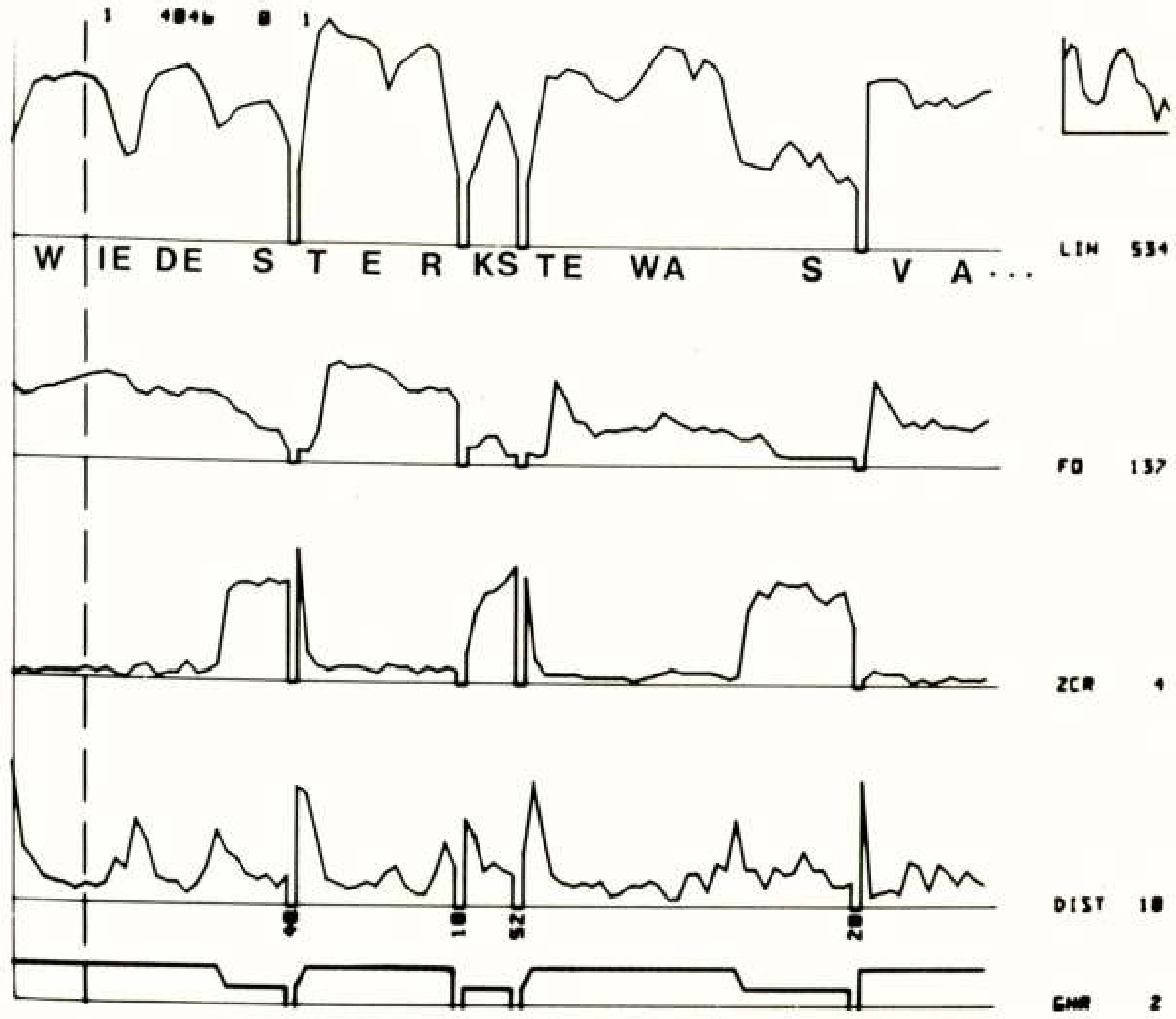


Fig. 2. Weergave van het tijdsafhankelijke verloop van een aantal parameters, zoals gefotografeerd van een computer display. Van boven naar beneden is achtereenvolgens weergegeven: het overall nivo, de grondfrekwentie, het aantal nuldoorgangen, de afstand tussen opeenvolgende spektra in de 17-dimensionale ruimte, en stemhebbend/stemloos. Tevens zijn op de betreffende plaatsen de pauzeduren in msek aangegeven. De stippellijn is een verplaatsbare pointer. In de rechterhelft van de display zijn de met dat tijdmoment overeenkomstige numerieke waarden van de diverse parameters aangegeven. Tenslotte is rechts bovenin het momentane bandfilterspektrum weergegeven. Het stuk tekst waarvan hier de analyseparameters zichtbaar zijn is: "Wie de sterkste was van ... ".

derzoek, waarbij we ook de toonhoogte vrijelijk wilden kunnen modificeren, minder geschikt. Wij werken dan ook met een vaste periodieke of ruisbron, voor het genereren van de stemhebbende en stemloze spraakklanken. Echter in plaats van de uitgangen van de filters te moduleren, introduceren wij de spektrale variatie door het bronsignaal zelf per filter te varieren. De synthetisator is een perifeer apparaat dat door middel van het op ons instituut ontwikkelde digitale input-output systeem vanuit de computer per 10 msek van nieuwe data wordt voorzien, waarna het apparaat zelfstandig de signaalgeneratie verzorgt.

De voor de synthese benodigde toonhoogte informatie wordt tijdens de analyse verkregen middels een keelmikrofoonsignaal. Door onder programmakontrôle de parameters nodig voor de synthese te modificeren kan de teruggegenereerde spraak op vele manieren worden beïnvloed. Zo kan een klein gedeelte worden geisoleerd en in detail uitgeluisterd, door herhaald genereren en/of uitrekken van de tijdschaal, zonder spektrale vervorming. De binnen- en tussenwoord pauzes kunnen worden benadrukt door ze te verlengen, of juist geheel over te slaan. De intonatie van de zin kan worden gewijzigd door modifikaties in het verloop van de toonhoogte aan te brengen.

Ook kan rekenkundig of eventueel in hardware een datareduktie op de basisgegevens worden toegepast, waarna hersynthese met een lagere bitrate kan worden gerealiseer. Tijdens de voordracht op 12 mei 1976 werden door middel van een bandopname een aantal van deze mogelijkheden gedemonstreerd.

Alhoewel flexibiliteit en eenvoudige modificeerbaarheid eerste vereisten van het systeem zijn, willen we toch ook iets zeggen over de spraakverstaanbaarheid van teruggegenereerde spraak. De spraakverstaanbaarheid is op de beproefde wijze gemeten via lijsten met eenlettergrepige nonsens woorden (fonetisch gebalanceerd = PB words), ingesproken door vijf verschillende sprekers, en beluisterd door vijf verschillende luisteraars. De gemiddelde PB-word verstaanbaarheid bedraagt 77,8%. Dit is voor onze toepassingen ruim voldoende en garandeert een voortreffelijke zinsverstaanbaarheid.

Wanneer datareduktie wordt toegepast neemt de woordverstaanbaarheid uiteraard af, en wordt 44,1% wanneer 3-dimensionale in plaats van de oorspronkelijke 17-dimensionale spektrale informatie wordt gebruikt.

Ons systeem biedt ook de mogelijkheid tot een zogenaamde pattern matching vocoder (Flanagan, 1972). Hierbij wordt gewerkt met een beperkt aantal diskrete spektra die gelabeled zijn, alleen deze nummers worden overgestuurd. In plaats van met diskrete spektra zou in ons geval met diskrete posities in een laag-dimensionale subruimte worden volstaan.

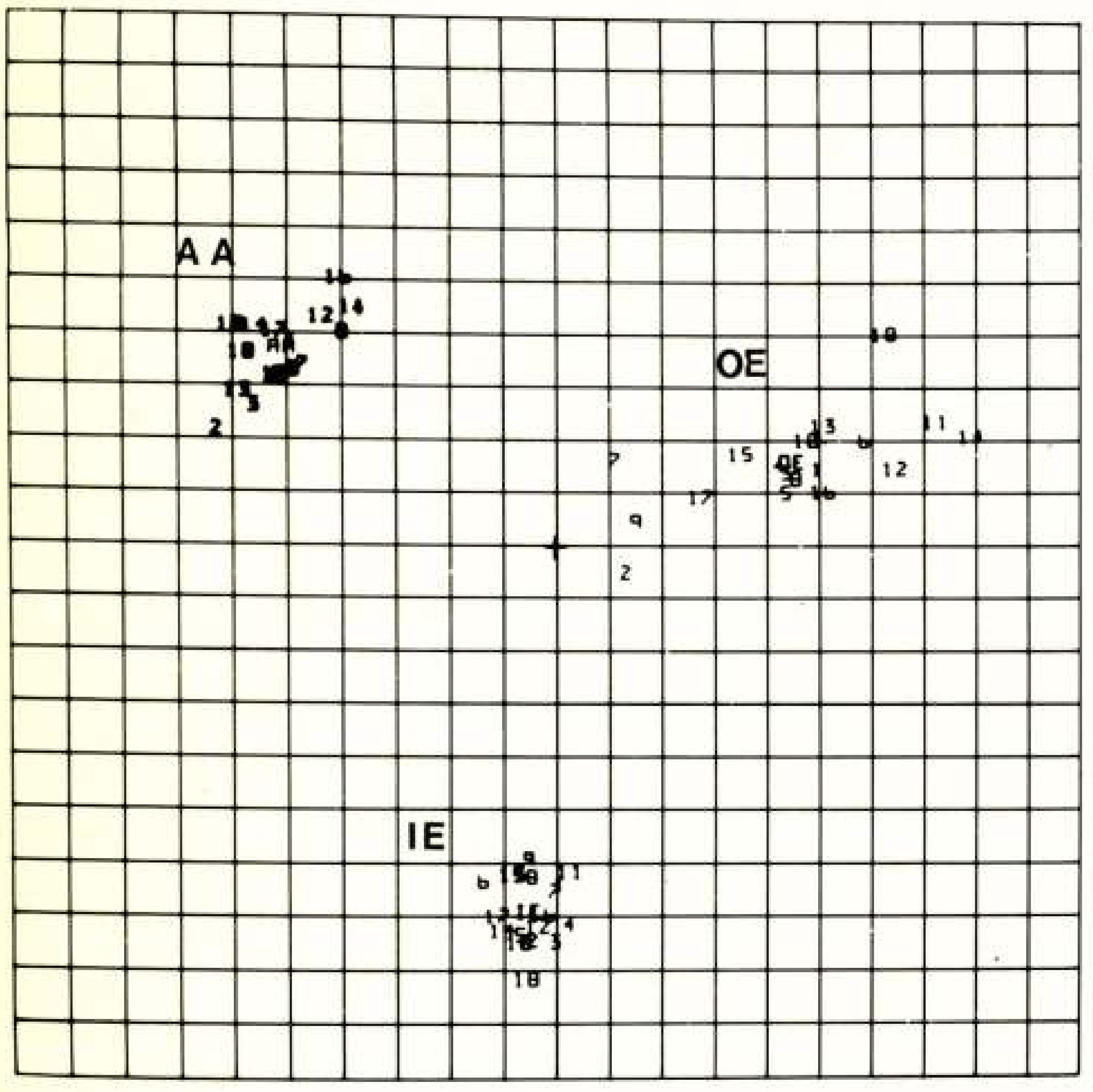


Fig. 3. Computer display van de gemiddelde posities van 18 klinkersegmenten geisoleerd uit eenlettergrepige woorden voor de klinkers *ie*, *oe*, en *aa*, weergegeven in de twee-dimensionale spektrale subruimte.

Tot slot een summiere beschrijving van twee typen onderzoek waarvoor het analyse-synthese systeem wordt gebruikt.

Onderzoek met het analyse-synthese systeem.

Teneinde enig inzicht te krijgen in de belangrijkheid van bepaalde frekwentiegebieden voor de perceptie van bepaalde klinkers is een benoemingseksperiment uitgevoerd met selektief gemaskeerde klinkers. Met behulp van de synthetisator werden hiertoe korte klinkers gegenereerd waarbij de spektrale informatie in één van de filters was vervangen door een band ruis. Als dit gebeurt in een frekwentiegebied dat niet erg bijdraagt tot de korrekte identificeerbaarheid van een bepaalde klinker, dan zal de maskering nauwelijks tot foutieve benoemingen leiden. Wanneer echter essentiele spektrale informatie wordt gemaskeerd dan zal dit leiden tot specifieke foutieve benoemingen (Pols, 1975).

Een heel ander type onderzoek betreft de fysische en perceptieve verschillen tussen klinkers in eenlettergrepige woorden. De variabelen die hierbij een rol spelen zijn de verschillnede Nederlandse klinkers, de verschillende sprekers, en de verschillende medeklinkeromgevingen waarin de klinkers kunnen worden geplaatst.

Het analyse-synthese systeem wordt gebruikt voor het analyseren en isoleren van de klinkersegmenten, en voor het hersynthetiseren van deze segmenten voor benoemingsproeven. Niet alleen de tweeklanken *ei*, *au* en *ui* blijken een sterk dynamisch verloop te hebben, maar ook klinkers als de *ee*, *oo* en *eu*, waarvan het verloop zich dan ook nog sterk wijzigt als deze klinkers voorafgaan aan de *r*.

Soesterberg, mei 1976

LITERATUUR

Flanagan, J.L. (1972). *Speech analysis synthesis and perception*, Springer Verlag, Berlin, 2nd edition.

v.d. Geer, J.P. (1967). *Inleiding in de multivariate analyse*, Van Loghum Slaterus, Arnhem.

Kuijk, K.E. en Fransen, N.V. "Reductie van spraakbandbreedte m.b.v. een vocoder, welke gebruik maakt van vooruitregeling", dit nummer.

Plomp, R. (1976). "Auditieve functies", Hoofdstuk 7,2 in *Handboek der Psychonomie*, Eds. Michon, J.A., Eijkman, F.G.J. en de Klerk, L.F.W., Van Loghum Slaterus, Deventer.

Pols, L.C.W. (1975). "Dominant spectral regions in vowel perception", Paper 347 of the 8th Congress of Phonetic Sciences, Leeds.

Povel, D.J.L. (1974). "Articulation correction of the deaf by means of visually displayed acoustic information", dissertatie K.U. Nijmegen.

Sluyter, R.J. (1976). "Een eenvoudige vocoder voor digitale spraakoverdracht", dit nummer.

Voordracht gehouden op 12 mei 1976 tijdens een gezamenlijke vergadering van het NERG (256ste werkvergadering), de Benelux-section IEEE, en het NAG, op het Instituut voor Zintuigfysiologie TNO, te Soesterberg.

U R S I

Het Symposium on Radio Waves and the Ionosphere gehouden te Lima, augustus 1975.

Commissie IV (magnetosfeer).

De talrijke onderzoekingen op het gebied van deze commissie betreffen het physische gedrag van het plasma in de magnetosfeer (dit is het gebied boven de ionosfeer alwaar de aanwezigheid van het aardmagneetveld wezenlijk is voor de physische structuur van de ruimte). De het langst bekende hieronder vallende verschijnselen zijn de "whistlers", vakterm voor de voortgeleiding van electromagnetische golven van zeer lage frequentie (tot in het hoorbare gebied) langs de krachtlijnen van het aardmagneetveld. Deze geleiding wordt bevorderd door de vorming van "ducts", dit zijn ongeveer cilindrische buizen rondom bepaalde krachtlijnen waarbinnen de electronendichtheid het grootst is langs de as. Oorspronkelijk waren als bronnen van deze "whistlers" slechts onweershaarden in tropische gebieden bekend, terwijl hun ontvangst op aarde aan het einde van een magnetische krachtlijn het eerst in 1888 geconstateerd werd als storende geluidjes op een lange telefoonlijn die effectief was als ontvangstantenne voor zeer lage frequenties. Later bleek dat de voor "whistlers" karakteristieke voortgeleiding langs "ducts" zijn oorsprong even goed kan hebben in locale magnetosferische verstoringen die veroorzaakt worden door de wisselwerking van het ionosferische plasma met deeltjes van de daarop van boven invallende zonnewind. Recente op de conferentie in Lima behandelde onderzoekingen hebben vooral betrekking op kunstmatig (door injecties van stralen van geladen deeltjes) opgewekte whistlerverschijnselen. De resultaten daarvan verschaffen steeds meer inzicht in de physische verschijnselen die voortdurend in het magnetosferische plasma plaatsvinden, in het bijzonder in de wisselwerking aldaar tussen electromagnetische golven onderling en tussen dergelijke golven en geladen deeltjes. De combinatie van de vele aldus uitgevoerde experimenten leidt tot steeds uitvoeriger numerieke waarden van de ionospherische parameters zoals electronendichtheid, electronen en ionen temperaturen, botsings en recombinatie parameters, enz. Deze gegevens kunnen bovendien vergeleken worden met die afgeleid uit de waarnemingen van "incoherent scatting" (genoemd door Vesseur, op pagina 65 en 66). Een in het bijzonder in de laatste tijd verricht onderzoek houdt zich bezig met de aankomstrichting van op het aardoppervlak ontvangen whistlersignalen; deze laatste waarnemingen verschaffen nieuwe gegevens over de nog niet geheel begrepen wijze waarop een van de aarde afkomstige whistler het begin van een der bovengenoemde "ducts" bereikt, alsmede waarop deze na aankomst in het einde van deze duct het verderop gelegen ontvangpunt bereikt. Ook werden in Lima programma's voor het toekomstige onderzoek van al deze verschijnselen besproken.

Prof. dr. H. Bremmer
Technische Hogeschool Eindhoven

UIT HET NERG

LEDENMUTATIES

Voorgestelde leden

Ir. P.B. Hesdahl, Paduaweg 75, Den Dolder.
Ir. J.P. Kingma, Den Burghstraat 19, Voorburg Z.H.
Ir. J. Mendrik, Bosboom Toussaintplein 123, Delft.

Nieuwe leden

H.J.Middendorp, Witbreuksweg 383-312, Enschede.
Dipl. Ing. H.-G. Nussbaum, Jagersweg 4, Huizen N.H.
Ir. W.J.P.A. Verbeek, Ursulaland 147, Den Haag.
R.G. van Vliet, Jan Bollandlaan 21, Eindhoven.
Ir. J.R. Westerveld, Van Adrichemstraat 191, Delft.
Ing. K. Zeehuisen, Donizettihof 23, Alphen aan de Rijn.

Nieuwe adressen van leden

Ir. C.A. Bruijnis, Zuiderdreef 51, Nieuw Vennep.
Dr. ir. H.C.A. van Duuren, Raaphorstlaan 25, flat 330, Wassenaar.
Ir. T.L. Jap, Noordeinde 3, Delft.
Ir. C. Kamminga, Nieuwe Plantage 47, Delft.
Ir. L.P. Ligthart, Valenberg 63, Zoetermeer.
Ir. P.K. Reek, Paradijspoort 51, Delft.
Ir. J. van Staalduinen, Gebr. de Wittplantsoen 17, Voorschoten.
Ir. J.N. Vles, Nieuwe Vaart 10, Vleuten.

ESSCIRC 76

2 ND EUROPEAN SOLID STATE CIRCUITS CONFERENCE
SEPTEMBER 21 - 24, 1976
UNIVERSITE PAUL SABATIER, TOULOUSE - FRANCE

PRELIMINARY PROGRAMME
REGISTRATIONS

LAAS CNRS 7, avenue du Colonel Roche, 31400 Toulouse, France

Cosponsors

Associazione Elettrotechnica ed Elettronica Italiana, European Physical Society, Convention of National Societies of Electrical Engineers of Western Europe, The Institution of Electrical Engineers, The Nachrichtentechnische Gesselschaft im Verband Deutscher Elektrotechniker, The Nederlands Elektronica-En Radiogenootschap, the « Société des électriciens, des électroniciens et des radioélectriciens », the « Société française de Physique ».

Professional and financial contributions

Centre National d'Etudes Spatiales, Centre National d'Etudes des Télécommunications, Centre National de la Recherche Scientifique, Direction des Industries Electroniques et de l'Informatique, Délégation Générale à la Recherche Scientifique et Technique, Institut de Recherche en Informatique et Automatique, Laboratoire d'Automatique et d'Analyse des Systèmes du C.N.R.S., Laboratoire d'Electronique et des Techniques d'Information CEA.

Programme Committee

Chairman	: J. LACOUR (France)	Organizing Committee
Vice Chairman	: H. MARTINOT (France)	
Secretary	: D. ESTEVE (France)	
Members	: J. BOREL (France)	
	M. BRUNET (France)	
	J. CACHEUX (France)	
	J. S. CAVE (Great Britain)	
	E. CONSTANT (France)	
	P. DELIGNE (France)	
	C. DURANTE (France)	
	W. ENGL (Germany)	
	G. GRUNBERG (France)	
	P. JESPERS (Netherlands)	
	P. N. ROBSON (Great Britain)	
	K. STEIN (Germany)	
	V. SVELTO (Italy)	
	N. DE TROYE (Netherlands)	
	E. VITTOZ (Switzerland)	

Corresponding members :

M. NAGATA (Japan)
K. ZAININGER (U.S.A.)

GENERAL INFORMATION

1. General Organization

— All sessions will be held in the Conference rooms of the « Université Paul Sabatier », 118, Route de Narbonne, 31400 Toulouse.
— A bus service will operate between hotels and the Conference Place.
— The reception service, where participants will obtain their documents, will be open from Monday 20 September from 14 h to 18 h at the Entrance Hall of the « Université Paul Sabatier », and Tuesday from 8 h 30.
— The invited papers will preceed the sessions, certain of the sessions being parallel. English will be the official language of the Conference.
Opening of the Conference : September 21, 1976 at 9 h 30.

2. Documentation

— A set summaries of the contributed papers will be given to the participants at the beginning of the Conference.
— The « Invited Papers » will be published as a separate volume of the conference series by the « Journal de Physique ». A copy will be sent in the Spring of 1977 to all registered participants. Extra copies will be available from the « Journal de Physique ».

3. Social events

Social events, which will be planned by the Conference organizers, will take place throughout Conference : A reception on the evening of Tuesday 21 by the Mayor of Toulouse, an afternoon visit of Toulouse on Wednesday 22, and a Conference Banquet on Friday at noon.
A ladies' programme will be arranged on request.

4. Fees

Registration fees for each participant have been fixed at 550 FF.

The registration fee includes :

a set of summaries
the book of the invited papers, to be published in the Spring of 1977
lunch for 3 days
refreshments
a bus service operating between hotels and the Conference place
social events.

For accompanying persons a contribution of 80 FF will be requested for the Conference Banquet.
For students, who can justify their status, there will be a special fee of 250 FF but which will not include participation in the social events, nor will they receive the book of the invited papers.

5. Registrations

Those wishing to participate are invited to complete the enclosed registration form and return it to ESSCIRC 76 Secretariate.

6. Hotel accomodations

Your hotel reservation will be made through the « Syndicat d'Initiative » of Toulouse.
For details, please refer to the enclosure.
Post congress tours will also be organized after the symposium by the « Syndicat d'Initiative ».

6. Secretariate

The secretariate for the conference will be located as follows :
until September 21, 1976

LAAS CNRS
7, avenue du Colonel-Roche
31400 TOULOUSE (France)
Tel. : (61) 53.11.77 Telex : LAASTLSE 520930

from September 21 to 24 inclusive

Université Paul Sabatier
118, route de Narbonne
31400 TOULOUSE (France)
Tel. : (61) 52.15.95.

LATE NEWS PAPERS

The closing date for regular contributions was May 5. However late news papers may be submitted to the Program Committee. These should be no longer than five minutes presentation time. Abstracts of no more than 300 words typed, double space, on one side of an A_4 should be submitted in duplicate to the general secretary of the Conference : D. Estève, LAAS, CNRS. Notification of acceptance or rejection cannot be given to the contributors before conference time.

PRELIMINARY PROGRAM

TUESDAY SEPTEMBER 21

9.30	Conference Opening
10.30	**Invited Paper**: LSI in Watches - D.E. VITTOZ - Centre Electronique Horloger - Neuchâtel
11.30	Break

SESSION I: TELECOMMUNICATION ELECTRONICS

11.40 A-1-1	High density monolithic bipolar and MOS crosspoint arrays for private branch telephone exchanges E. BACHLE - H. CLAUSS - J. DANGEL - G. KOHLBACHER - K. SCHLUTER - H. SCHUSSLER - H.J. WULF - AEG TELEFUNKEN
12.00 A-1-2	An I²L two-tone dialling circuit for telephony J.C. KAIRE - P. GUETIN - RTC La Radiotechnique Compelec - Caen - D.J.G. JANSSEN - PHILIPS Gloeilampenfabrieken - Eindhoven
12.20 A-1-3	Integrable electronic gyrators with high efficiency E. HOCHMAIR - Institut für Physikalische Elektronik Technische - Wien
12.40	**LATE NEWS PAPERS**
13.00 - 14.00	Lunch
14.00	**Invited Paper**: Semiconductor technologies with reduced dimensions K.U. STEIN - D. WITMAN - SIEMENS - Munich Research Laboratories

SESSION II: INTEGRATED ELECTRONICS

15.00 A-2-1	New analog CMOS IC's based on weak inversion operation E. VITTOZ - J. FELLRATH - Centre Electronique Horloger S.A. - Neuchâtel
15.20 A-2-2	A low drift MOSFET operational amplifier ARZ R. POUJOIS - J.M. ITTEL - J. BOREL - CEA CENG LETI - Grenoble
15.40 A-2-3	A differential MOS-bipolar voltage buffer with virtually zero input capacitance H. DE MAN - R. VAN PARIJS - R. CUPPENS - Katholieke Universiteit Leuven - Laboratorium voor fysica en elektronica van de halfgeleiders - Heverlee
16.00	Break
16.20 A-2-4	A New circuit technique in voltage and temperature compensated emitter-coupled logic W. WILHELM - K.R. SCHON - SIEMENS - Munich
16.40 A-2-5	Digital tuning of television receivers using nonvolatile memory T.K. DAS - P.H. BOYCE - G.J. FLETCHER - Plessey Semiconductors - Swindon
17.00 A-2-6	Monolithic nullor J.H. HUIJSING - J. DE KORTE - Delft University of Technology
17.20 A-2-7	Quad JFET wideband operational amplifier IC D.L. CAVE - M.P. DAVIS - Motorola - USA

SESSION III: OPTOELECTRONICS

16.15 B-3-1	An improved monolithic charge-sensing circuit for analog arrays. P. JESPERS - C. JUSSERET - Y. LEDUC - D. COLARD - M. BONNEJONNE - Université Catholique de Louvain - Laboratoire de Microélectronique
16.35 B-3-2	Investigations on high-resolution imaging with three-terminal CID's R. KOCH - SIEMENS - Munich
16.55 B-3-3	Optimal use of junction isolation for photodetection in monolithic integrated optoelectronic circuits K. BERCHTOLD - S. DERMITZAKIS - J. SURI - AEG TELEFUNKEN - Heilbronn
17.15 B-3-4	Optical character recognition device W.D. RYAN - D.H. CAMPBELL - The Queen's University of Belfast
18.30	Reception by the Mayor of Toulouse

WEDNESDAY SEPTEMBER 22

9.00	**Invited Paper**: From where to where in logic arrays H. FLEISHER - IBM Corporation - New York

SESSION IV: MEMORIES

10.00 A-4-1	Design of a very large reconfigurable memory J.M. AYACHE - ENSIMA - Grenoble
10.20 A-4-2	Simple Al-gate-technology yields high bit-density G. MEUSBURGER - H. KELLER - SIEMENS - Munich
10.40	Break
11.00 A-4-3	A 16-kilobit CCD memory J.N. GOODING - Allen Cleark Research Centre
11.20 A-4-4	A 16 K MOS RAM with double-polysilicon technology R.W. MITTERER - B.F. REHN - SIEMENS - Munich
11.40 A-4-5	1024-bit fully decoded MNOS non-volatile memory D. BOSTOCK - Allen Clarck Research Centre
12.00 A-4-6	A 1024 bits RMM memory with TTL compatibility E. MACKOWIAK - V. LE GOASCOZ - CEA CENG LETI - Grenoble
12.20 A-4-7	Project on an electrically reprogrammable fast read-out memory using bipolar tetrode transistors J.P. DOM - Ph. ROUX - J.L. AUCOUTURIER - SEL Université de Bordeaux - M. DEPEY - SESCOSEM - Saint-Egrève
12.40	**LATE NEWS PAPERS**

SESSION V: MODELING

11.00 B-5-1	Objectives for integrated circuit tolerance design P.R. ADBY - University of London King's College
11.20 B-5-2	Energy per logic operation - A figure of merit for IC'S R. MULLER - H.J. PFLEIDERER - K.U. STEIN - SIEMENS - Munich
11.40 B-5-3	A high performance LSI circuit design strategy S. HOLLOCK - Allen Clark Research Centre - Towcester
12.00 B-5-4	Thermal coupling phenomena in IC'S: Models for analysis and synthesis for circuits based on thermal coupling V. SZEKELY - K. TARNAY - Technical University Budapest
12.20 B-5-5	A new method for the determination of transfer inefficiency in MOS charge coupled devices B. MEUSEMANN - E. FROESCHLE - Technische Hochschule Aachen
13.00 - 14.00	Lunch

THURSDAY SEPTEMBER 23

9.00	**Invited Paper**: Present and future of bipolar technology V. BLATT - Allen Clark Research Centre
10.00	**Invited Paper**: Advanced MOSFET technologies: A review J. BOREL - LETI
11.00	Break

SESSION VI: INTEGRATED INJECTIONS LOGICS

11.15 A-6-1	Modeling of I²L injection logic structures J.P. BAILBE - A. MARTY - LAAS CNRS - Toulouse
11.35 A-6-2	Stacked I²L circuit M. NAGATA - K. KANEKO - T. OKABE - Central Research Laboratory of Hitachi
11.55 A-6-3	Twin-transistor-injection-logic (T²IL) G. NUZILLAT - T. PHAM NGU - H. HELIOUI - THOMSON CSF - Orsay
12.15 A-6-4	Schottky I²L (substrate FED logic) - An optimum form of I²L P.S. WALSH - G.W. SUMERLING - Allen Clark Research Centre - Towcester
12.35 A-6-5	Schottky collector integrated injection logic S. BLACKSTONE - R. MERTENS - Université Catholique de Louvain
13.00 - 14.00	Lunch
14.00	**Invited Paper**: Silicon on Sapphire technology Y. NISHI - Toshiba - Japan

SESSION VII: MOS CIRCUITS

15.00 A-7-1	Low-power D-MOS logic M. DECLERCQ - Th. LAURENT - Université Catholique de Louvain
15.20 A-7-2	N-channel-silicongate-technology for LSI-applications J. DANGEL - W. LANGHEINRICH - E. STEIN - A. STURMER - AEG TELEFUNKEN - Ulm
15.40	Break
16.00 A-7-3	Circuit implications of the P.S.D.-MOST process Ir. L. SPAANENBURG - Twente University of Technology
16.20 A-7-4	CO 8: An eight channel multiplexer on silicon on sapphire J.P. SUAT - G. MORIN - CEA CENG LETI - Grenoble
16.40 A-7-5	Resistance C-MOS circuits H. OGUEY - E. VITTOZ - Centre Electronique Horloger S.A. - Neuchâtel
17.00 A-7-6	A C-MOS/SOS technology for radiation tolerant devices J. PINEL - J.M. CHARLOT - J.P. SUAT - J. BOREL - CEA

SESSION VIII: DESIGN AND TESTING

16.00 B-8-1	Fault detection problems and their impact on the design of digital circuits M. DIAZ - M. COURVOISIER - J.C. GEFFROY - LAAS CNRS - Toulouse
16.20 B-8-2	A standardized method reduces design time of C-MOS integrated circuits and enables automatic checking J.C. BERTAILS - J. ZIRPHILE - SESCOSEM
16.40 B-8-3	CAD models of MOSFET transistors and their parameter acquisition on an automatic system B. BAYLAC - G. MERCKEL - CEA CENG LETI - Grenoble
17.00 B-8-4	Laser testing of integrated circuits J.G. SMITH - H.E. OLDHAM - University of Southampton
17.20	**LATE NEWS PAPERS**

FRIDAY SEPTEMBER 24

9.00	**Invited Paper**: High speed Ga As circuits R. VEILEX - Laboratoire d'Electronique et de Physique Appliquée
10.00	Break

SESSION IX: HIGH SPEED CIRCUITS

10.20 A-9-1	Ga As integrated circuits for high speed logic D. BOCCON-GIBOD - G. DURAND - L.E.P.
10.40 A-9-2	Ga As MESFETs applications H.J. KLEIN - W. FILENSKY - Technical University Aachen
11.00 A-9-3	2 GHz logic based on a 5 GHz fT process M. KITCHIN - P.S. WALSH - P.J. WARD - Allen Clark Resarch Centre - Towcester
11.20 A-9-4	Travelling wave dividers up to 1.4 GHz D. GRENIER - RTC - Caen D. KASPERKOVITZ - PHILIPS Researches Laboratories - Eindhoven
11.40	**LATE NEWS PAPERS**

SESSION X: COMPUTER AIDED DESIGN

10.20 B-10-1	Astec 3 - A user oriented circuit analysis program M.H. HEYDEMANN - CEA - Montrouge
10.40 B-10-2	Episode - A set of tools oriented to logic integrated circuit design verification and testing C. CHICOIX - L.A.M. J. THUEL - LETI R. TULLOUE - THOMSON CSF
11.00 B-10-3	CALMOS: a computer aided layout program for MOS/LSI H. BEKE - W. SANSEN - R. VAN OVERSTRAETEN - Katholieke Universiteit Leuven
11.20 B-10-4	CIRCEC: A cost-effective computer program for circuit design D. HERRERIAS - A. LABARTHE - G. GUERIN - THOMSON CSF
11.40 B-10-5	Transient of nonlinear periodic transient (TNPT) - A new simulation feature H.P. STROHBAND - R. LAUR - Institut für Theoretische Elektrotechnik - Aachen
12.30	Conference Banquet

Tijdschrift van het Nederlands Elektronica- en Radiogenootschap

Inhoud

deel 41 - nr. 3 - 1976

druk: HET ZUIDEN Eindhoven